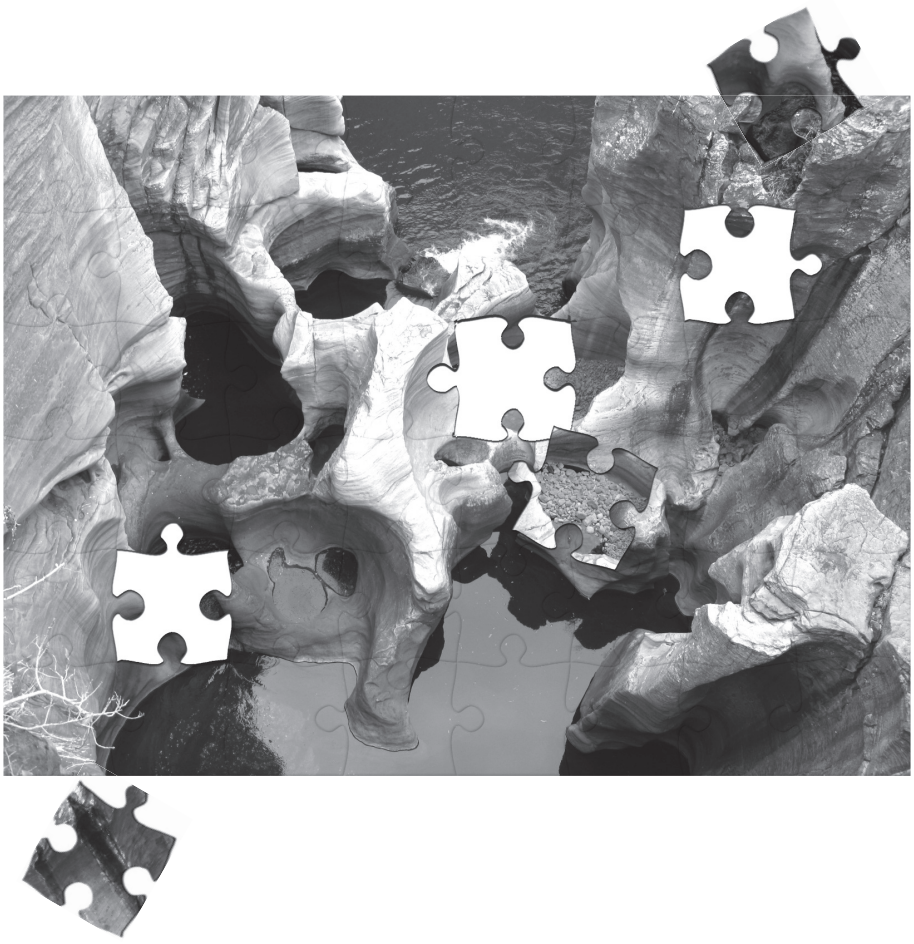
Heterogenomics of autoimmunity towards personalized medicine

van Baarsen, E.G.M.

2009

document version
Publisher's PDF, also known as Version of record

Heterogenomics of autoimmunity towards personalized medicine



Lisa G.M. van Baarsen

The studies described in this thesis were performed at the department of Molecular Cell Biology & Immunology and the department of Pathology of the VU University Medical Center, Amsterdam, Netherlands.

This project was financially supported by a V-ICI grant from the VUMC

Printing of this thesis was financially supported by:
J.E. Jurriaanse Stichting
The Dutch Arthritis Foundation
Stichting MS Research
Abbott B.V
Greiner Bio-One
Tebu-Bio

Cover photo: "The heterogeneity of potholes", Bourkes Luck Potholes, Blyde River Canyon, Mpumalanga, South Africa.

Cover design and photo: Lisa G.M. van Baarsen

Layout: Jo-Ann Snel, Op Honk Design - Uitgeverij Boekenbent

Printed by: Wöhrmann Print Service

ISBN: 978-90-8570-442-3

VRIJE UNIVERSITEIT

Heterogenomics of autoimmunity towards personalized medicine

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. L.M. Bouter,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Geneeskunde
op donderdag 1 oktober 2009 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Elisabeth Geertruida Maria van Baarsen

geboren te Volendam

promotor:	prof. dr. C.L. Verweij
copromotor:	dr. C.T.M. van der Pouw Kraan

"Wetenschap definieert niet maar dringt tot de kern door, trekt geen scheidslijnen maar schept overgangen, trekt geen dogma's maar ontwikkelingen. Wetenschap kan niets bewijzen, maar alleen door het wegstrepen van variabelen een zo duidelijk mogelijk beeld schetsen." *De Zee, Frank Schätzing*

Ter nagedachtenis aan Marcel van Baarsen
7 mei 1977 – 28 juni 2009

Beoordelingscommissie:

Prof. Dr. R. Mebius
Dr. L. Van 't Veer
Prof. Dr. B.A.C. Dijkmans
Prof. Dr. P.P. Tak
Prof. Dr. C.H. Polman
Prof. Dr. L.A. Aarden
Prof. Dr. G.A. Meijer

Table of contents

Chapter 1. Introduction

1

Parts of this chapter are adapted from:
A genomic view of subtypes in rheumatoid arthritis: Towards personalized medicine van Baarsen E.G.M., van der Pouw-Kraan C.T.M, Verweij C.L.
Future Rheumatology June 2006 Vol. 1, No. 3, Pages 311-322

Gene expression profiling in rheumatology
Tineke C.M.T. van der Pouw Kraan, Lisa G.M. van Baarsen, Francois Rustenburg, Belinda Baltus, Mike Fero and Cornelis L. Verweij
Book chapter In: Methods in Molecular Medicine (ed. Andrew Cope), Arthritis Research; Methods Mol Med. 2007 Vol.136, Pages 305-327

1

1.1 Autoimmunity

Under normal conditions, immune cells that recognize self antigens are eliminated by apoptosis upon encounter with self-antigen during ontogeny or alternatively, are made unresponsive to self through anergy, receptor editing, or lack of expansion or maturation within or outside the germinal centers (1;2). However, when the response of the immune cell is not managed correctly after contacting a self antigen, an immune response against the self antigen may be induced. This undesired autoimmune response can lead to autoimmune diseases leading to by chronic inflammation and tissue damage. These diseases may be characterized by the presence of self-antigen specific T- cells and/or specific autoantibodies which are indicative for the autoimmune response.

The incidence of most individual autoimmune diseases is low, but in composite, autoimmune diseases affect approximately 5% of the population in Western countries (3). These chronic inflammatory autoimmune diseases are distinguished on the basis of clinical manifestations in organ-specific and systemic inflammatory diseases. Autoimmune diseases like multiple sclerosis (MS) and insulin-dependent diabetes mellitus (IDDM) are referred as organ-specific whereas rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are systemic autoimmune diseases affecting several different compartments of the body. It is believed that common biological processes may be involved in development of different autoimmune diseases.

It is unknown what triggers autoimmunity but both environmental and genetic factors, especially HLA alleles, are clearly important. Until recently, most of the treatment strategies for these diseases were dependent on non-specific immunosuppressive approaches which can cause serious side effects. However, recent progress in understanding the pathogenic mechanisms that contribute to autoimmune diseases has enabled the development of targeted therapies using biologicals, which are (parts of) natural occurring proteins such as antibodies against molecules such cytokines.

1.2 Key cytokines involved in immunity

Autoimmune diseases are characterized by local tissue injury caused by the persistent presence of activated infiltrating cells which produce inflammatory mediators such as cytokines and chemokines leading to tissue damage and the recruitment of new infiltrating cells. To date, many chemokines and cytokines are described each with specific pro- or anti-inflammatory roles in normal homeostasis and immunity. Activated cells can secrete a large variety of cytokines with different local or distal effects. Cytokines that are known to play a major role in chronic inflammatory diseases are TNF and type I interferons (IFN).

1.2.1 TNF

TNF is a pleiotropic cytokine mediating a wide variety of biological activities which are extensively studied in both experimental animal models as well as in humans (reviewed in (4)). TNF has an important role in normal physiology as well as in response to inflammation, but inappropriate or excessive levels can induce inflammation and organ injury. TNF exists in both transmembrane and soluble forms but it is not well understood which one contributes most the inflammatory response (5). Soluble TNF (sTNF) preferentially binds to TNFR1 (6) while transmembrane TNF (mTNF) primarily signals through TNFR2 (7). Whereas TNFR1 mediated signalling is well known for its pro-inflammatory and apoptosis inducing capacities, signalling through TNFR2 is quit less understood. Moreover, mTNF by itself may lead to intracellular signalling (8) making the TNF induced signalling program even more complex. TNF plays a pathogenic role in several immune-mediated inflammatory diseases such as RA, Chrohn's disease, psoriasis and others. Therefore, therapies aimed at inhibiting the function of TNF have been developed and are successfully used as a key treatment strategy in RA and Crohn's disease patients to inhibit further disease progression (9).

1.2.2 Type I IFNs

Type I IFNs were the first cytokines described by Isaacs and Lindenmann (10) and named because of their role in interfering with virus replication. The family of type I interferons (IFN) are well-known for their anti-viral capacities and consists of multiple subtypes of IFNα, a single IFNβ and some less characterized family members, such as IFNε, IFNκ and IFNω (11). Whereas in peripheral blood type I IFNβ is mainly produced by plasmacytoid DCs (pDC) (12), in peripheral tissues stromal cells like fibroblasts can produce high amounts of IFNβ. However all activated nucleated cells are capable of producing type I IFN. After secretion, type I IFN can bind to its receptors IFNAR1 and IFNAR2 which are coupled to tyrosine kinases JAK1 (Janus kinase 1) and TYK2 (Tyrosine kinase 2) which after phosphorylation activate a signalling cascade leading to phosphorylation of STAT (signal-transducer and activator of transcription) proteins and subsequent transcription of several IFN-induced genes. The type I IFN response program induces resistance to viral replication, increases MHC class I expression and antigen presentation and activates NK cells to kill virus-infected cells. Moreover, type I IFNs are known to induce differentiation of monocytes into professional antigen presenting DCs (13;14) resulting in an amplification of the immune response towards antigens.
Recent studies have shown a major role for type I IFN and IFN-induced genes in SLE (15;16). Conversely, type I IFN administered to MS patients is one of the most successful therapies to prevent further disease progression.

1

1.3 Opposing roles of cytokines in different autoimmune diseases

Experimental mouse models of SLE have demonstrated a direct relationship between type I IFN and systemic humoral autoimmunity (17;18). SLE is characterized by a loss of tolerance against nuclear antigens leading to the formation of anti-nuclear antibodies (ANA) in >95% of SLE patients. The secretion of these autoantibodies leads to the generation of nucleic acid-containing immune complexes that are able to stimulate B cells and pDCs to increase the production of type I IFN. The increased bioavailability of type I IFN is thought to contribute to breakdown of peripheral tolerance through "inappropriate" activation of immature DCs resulting in an immune response against self antigens (reviewed in (19)). Whereas in SLE type I IFN is associated with disease severity (15) and autoantibody production (20), it has beneficial effects in MS , an autoimmune disease of the central nervous system (CNS) (21). Furthermore, IFNβ is reported to inhibit collagen-induced arthritis in mice (22), although so far clinical trials in RA patients failed (23). These studies show that type I IFNs can have different immunomodulatory functions depending on the type of autoimmune disease.

Another example of such an opposing cytokine function is the role of TNF in autoimmune diseases. Whereas in RA anti-TNF treatment is beneficial for a large part of the patients, in MS this drug is associated increased MRI activity (24). Moreover, TNF antagonists may induce demyelination (25), reactivate tuberculosis infection (26) and induce symptoms of reversible SLE (27-30). Since SLE is associated with increased expression of type I IFN (15) and anti-TNF treatment may induce symptoms of SLE (31), it was suggested that TNF may regulate type I IFN production. Indeed, pDCs re-exposed to viruses showed an enhanced production of type I IFN in the presence of TNF antagonists. Accordingly, systemic onset juvenile idiopathic arthritis (SOJIA) patients treated with anti-TNF displayed an increased expression of type I IFN induced genes in peripheral blood cells compared to untreated patients (32). Conversely, type I IFN may regulate the production of TNF (33;34).

These opposing effects of type I IFN and TNF in autoimmunity are highly intriguing and suggest that the in vivo interplay between inflammatory cytokines may determine the balance between protective immunity and detrimental effects (Figure 1) (35).

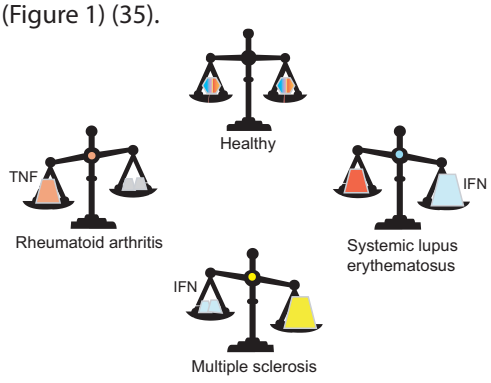


Figure 1. Cytokine balance in immunity

Cytokines can have opposing functions in different autoimmune diseases. E.g. while type I IFN is associated with disease severity in SLE, in MS it exerts ameliorative effects. In addition, TNF blockade is beneficial in RA patients while in MS patients it may among others induce demyelination.

1.4. Rheumatoid arthritis and multiple sclerosis

RA and MS are two prototypic autoimmune diseases driven by different cytokines. The most beneficial therapy in RA is TNF blockade, while this therapy has shown to result in adverse effects when applied in MS. Instead, therapy with type I IFN is the therapy of choice in MS. Therefore, these diseases may reflect two different ends of the spectrum of autoimmunity.

1.4.1 Rheumatoid arthritis

Rheumatoid arthritis (RA) is a complex, chronic, inflammatory autoimmune disease affecting mostly the joints. About 1% of the population is affected with a higher prevalence rate in females. The disease can appear at any age but is most common in people between 40 and 70 years of age. Much insight into the pathogenesis of disease has been gained during the last decade and although this has lead to the development of several therapeutics aimed at inhibiting disease progression, the disease can still not be cured nor prevented.
The predominant symptoms of RA are pain, stiffness and swelling of peripheral joints but it is a systemic disease with extra-articular involvement of other organ systems such as the skin and heart. Because of the broad range in clinical presentations it can be difficult to ascertain the diagnosis of RA (36). Diagnosis of RA is performed according to the 1987 revised ACR criteria (37;38) (Table 1).

Table 1. 1987 Criteria for the Classification of Acute Arthritis of Rheumatoid Arthritis**

Criterion	Definition
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement
2. Arthritis of 3 or more joint areas	At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints
3. Arthritis of hand joints	At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry)
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxta-articular regions, observed by a physician
6. Serum rheumatoid factor	Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in <5% of normal control subjects
7. Radiographic changes	Radiographic changes typical of rheumatoid arthritis on poster anterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify)

* For classification purposes, a patient shall be said to have rheumatoid arthritis if he/she has satisfied at least 4 or these 7 criteria. Criteria 1 through 4 must have been present for at least 6 weeks. Patients with 2 clinical diagnoses are not excluded. Designation as classic, definite, or probable rheumatoid arthritis is not to be made (37)

1.4.1.1 Preclinical phase of RA

The aetiology of RA is still unknown but both environmental as genetic factors are clearly important. Already early in disease the synovial joints are attacked by infiltrating immune cells (39) leading to synovitis and breaking down of cartilage and bone which can ultimately lead to joint destruction. Therefore, it is of high importance to diagnose disease as early as possible so treatment can be started to prevent malformation of the joints.

In RA the presence of specific autoantibodies against antigens containing one or more citrulline residues, the so-called anti-cyclic citrulline protein antibodies (ACPA), and rheumatoid factor (RF) are instrumental in the classifying diagnosis (38;40). Several studies have documented that RF and/or ACPA production precedes the development of RA. Nielen and colleagues (41) and Rantapää-Dahlquist and colleagues (42) analysed the levels of RF and ACPA in dated serum samples from RA patients who were former blood donors. They reported that RF and ACPA were already present up to 15 years before the appearance of the first clinical signs of arthritis. From these studies it is clear that production of these autoantibodies is an early process and that their appearance is predictive for the development of RA. It has been increasingly hypothesized that autoantibodies like ACPA are pathogenic and involved in the development of RA (43-45). Hence, the appearance of autoantibodies represents a subtle change in the immune system that probably reflects an early sign of pathogenesis which manifest itself without clinical symptoms. However, since not all RA patients are ACPA and/or RF positive and not all ACPA and/or RF positive individuals will develop RA, either additional factors are needed to result in a disturbed inflammatory immune response ultimately leading to RA or some individuals may have a protective immune profile which suppresses disease development despite the presence of autoantibodies. Hueber et al. suggest that anti-citrulline autoantibodies exhibit epitope spreading in the further preclinical and early phase of RA, suggestive of affinity maturation. Concomitantly, a time course-dependent cytokine signature evolves that predominantly associates with increased targeting of citrullinated autoantigens (46). Accordingly, additional research also revealed an increase in chemokine levels in patients positive for ACPA before disease onset (47). These findings might be an indication of an intimate co-regulation of a skewed immune response and autoantibodies as early processes in the development of RA. These initial immunological changes predispose to further progression of disease with clinical symptoms that may develop in various clinical courses. However, the mechanism to explain the disease inducing process remains to be determined.

Identification of preclinical features of the immune system associated with future development of RA will increase our understanding of the biological basis of disease and may lead to the discovery of additional biomarkers to improve predictive power. Moreover, recognition of the preclinical phase allows a timely start of treatment with the ultimate goal of primary prevention.

1.4.1.2 Heterogeneity of RA

The clinical presentation of patients with RA may reveal striking heterogeneity with a spectrum ranging from mild cases to severe and erosive disease. Clearly, heterogeneity is already reflected in the classifying diagnosis for RA, which is based on the presence of four out of seven criteria (Table 1), and thus indicates that different sets of criteria are applied to classify "the same" disease (37;38).
The heterogeneous nature of RA is also reflected at the level of the distribution of lymphocytes in the rheumatoid synovium, which reveals a remarkable patient-specific organization level (48-50). In about a quarter of the RA patients cellular infiltrates in the synovial tissue show a high degree of organization resembling germinal center (GC)-like structures normally observed in secondary lymphoid organs. In the remainder of the patients these GC-like structures are absent in the affected synovial tissues. These patients show either a diffuse or an aggregated T- and B cell infiltrate (51).
Recent studies have demonstrated the use of ACPA as a valuable specific serologic marker to distinguish RA from other non-erosive types of arthritis (41;42;52-54). In addition, several studies have demonstrated the prognostic value of ACPA by its ability to predict erosiveness of developing RA (55-59). The predictive ability of ACPA is also studied in combination with other disease parameters. ACPA together with RF appeared the best prognosticator for the development of persisting RA (42;52;60). In addition, it was reported that a combination of ACPA with HLA class DRB1 antigens is strongly associated with future onset of RA, thereby predicting the future development of RA (53;61;62). Thus, ACPA appears to identify a subgroup of patients with an increased chance for developing persistent and erosive disease.
Hence, a combination of biomarkers that may include RF and ACPA could reflect different aspects of the disease process and therefore might be useful for evaluating prognosis in individual patients with early rheumatoid arthritis (61).
The variation in responsiveness to virtually any treatment modality in RA is consistent with the heterogeneous nature of the disease (63;64). The heterogeneity most likely has its origin in the multifactorial nature of the disease, whereby specific combinations of environmental factor(s) and a varying polygenic background are likely to influence not only susceptibility but also the severity and disease outcome. The relative contribution of the different effectors and/or disease pathways may vary between patients and, perhaps, between different stages of disease. Hence, the cumulative data provides evidence that distinct pathogenic mechanisms contribute to disease in RA.

1.4.1.3 TNF blocking therapy in RA

TNF antagonists are approved worldwide for the treatment of various rheumatic diseases. Clinical experience indicates that there are 'responders' as well as 'nonresponders', but clear criteria for such classification are still lacking. Despite the highly beneficial effects of TNF-blocking agents in suppressing disease, clear

efficacy appears to be limited to a subset of RA patients (63;64). In addition, the pharmacology and mechanism of action of TNF blockade is not fully understood. Because of the observations that only about a quarter of the patients show a significant clinical improvement (ACR70 criteria) and a significant proportion of patients do not respond at all to TNF blockade, additional effectors and/or pathways are thought to contribute to disease. Similar observations have been made for other therapies like treatment with CTLA4Ig, which blocks the interaction of CD80/86 on antigen presenting cells with CD28 on T cells, and B cell ablation therapy (65;66).
Given the destructive nature of the disease it would be highly desirable to predict in an early stage the most beneficial treatment for the subgroup of patients at risk. If we rely solely on clinical or radiological manifestations we will probably be responding too late in order to maximize protection. Unfortunately, criteria to make selections of patients for optimal treatment and research purposes are currently lacking. Such criteria would be highly beneficial for patient stratification in order to assign homogeneous groups of patients for genetic studies and to improve the likelihood to observe efficacy of treatment. Ultimately, this may lead to a personalised form of medicine, whereby a specific therapy will be applied that is best suited for an individual patient.

1.4.2 Multiple sclerosis

Multiple sclerosis (MS) is a chronic inflammatory disease of the CNS resulting in substantial disability through recurrent deficits of sensation and of motor, autonomic, and neurocognitive function (67;68). These neurological episodes may arise between the age of 20 and 40 and affects predominantly woman. Symptoms of the disease include fatigue, muscle weakness, loss of vision, memory loss and difficulties with balance and coordination. Since these symptoms are rather broad and indistinct it is difficult to make a definite diagnosis.
Diagnosis is established based on clinical evidence of dissemination of more than one lesion both in time and space as determined by e.g. magnetic resonance imaging (MRI) of the CNS as described by the McDonald criteria (69). Analysis of cerebral spinal fluid (CSF) can be helpful in diagnosing MS since the protein fraction often consists of oligoclonal bands. However, since oligoclonal bands are also found in other disorders it cannot be used as a single diagnostic test. From a pathological point of view the disease is characterized by demyelination, gliosis and varying degrees of axonal pathology probably caused by migrating inflammatory cells invading the CNS.
Despite the development of several disease modifying treatments, the beneficial effects are partial and most of the drugs induce serious side effects such as fever and higher susceptibility to infection. So far no curative drug is available thus there remains a need for more specific therapies aimed at inhibiting disease progression without compromising the physiological functions of the immune system.

1.4.2.1 Risk factors associated with MS

The aetiology of the MS is unclear, but both environmental factors as well as multiple genetic factors contribute to disease susceptibility (70). Whole genome scans revealed strong evidence that certain loci are linked to the occurrence of MS and evidence exists that the genetic associations differ between MS subgroups (71). Studies in monozygotic twins have shown a concordance rate of only 25% indicating a contribution of non-genetic environmental factors (72).
Epidemiological studies revealed that the incidence of MS correlates with latitude (73) and that migration from one geographic area to another seems to alter a person's risk of developing MS indicating that infectious triggers, especially during childhood, may have important roles in MS aetiology. The most favoured hypothesis is that an abnormal immune response to an antigen causes the chronic inflammatory demyelinating disease. Several findings point into the direction that viruses or bacteria are causative in MS, although no single perpetrator has been identified (74-77). Anti-EBV antibodies are elevated in patients with MS compared to the general population (78) and herpes simplex virus is of interest because of its tendency of causing latent and recurrent infections. Moreover, it has been suggested that MS patients often relapse after a viral or bacterial infection (79). Furthermore, in animal models demyelination can be induced through peripheral immunization with a virus containing vaccine (80).

1.4.2.2 Heterogeneity of MS

The ever-growing body of both animal and human studies has led to the appreciation that MS is a heterogeneous disorder which is reflected in the variability of the disease at several levels. Clinically the disease can be divided into two major forms; most patients suffer from the relapsing remitting type of MS (RRMS) whereas 10-15% of the patients are diagnosed with the primary progressive form of MS (PPMS). In RRMS the episodes (relapses) of loss of motor and/or sensory function are followed by periods of (almost) complete recovery (remission). After approximately ten years of RRMS the disease switches to the secondary progressive phase of MS (SPMS) characterized by periods of symptoms that gradually accumulate without complete remission. Contrary to RRMS, PPMS mostly affects males, occurs later in life and from the beginning the disease symptoms gradually increase continuously.
Even within these predefined disease subtypes patients have a marked and largely unpredictable variability in disease severity, symptom spectrum, and treatment response. The latter is exemplified by the poor response to treatment in a significant number of RRMS patients using recombinant IFNβ treatment (discussed in next section). In addition, MRI, pathological, clinical, and epidemiological data indicate that the pathogenic mechanisms differ between MS patients (81;82). The molecular mechanisms behind this heterogeneity are only partly known.
Thus, it is becoming clear that, i. MS is a complex multi-factorial disease, ii. Multiple pathological patterns exist, and iii. These together accounts for differences in

clinical presentation, MRI phenotypes and treatment responses. This complexity hampers detailed insight in the pathogenesis of the disease.

1.4.2.3 IFNβ treatment in MS

IFNs were the first agents to show clinical efficacy in the most common form of MS, RRMS. Currently, a prolonged course of IFN treatment is still the best available therapy for RRMS. There are three preparations available for therapy: recombinant IFNβ1b produced in Escherichia coli ((Betaseron®/Betaferon® Chiron, Emeryville CA and Schering AG, Berlin, Germany) which was the first IFN approved for use in MS, and two types of recombinant IFNβ1a derived from Chinese hamster ovary (CHO) cells (Avonex®, Biogen Inc., Cambridge, MA, and Rebif®, Ares-Serono S.A., Geneva, Switzerland).

IFNß treatment reduces clinical relapses, has an ameliorating effect on brain disease activity, and possibly slows down progression of disability. In addition IFNβhas been found to suppress T cell proliferation, reduce T cell migration from the systemic circulation into the CNS, and alter the T cell cytokine secretion repertoire from a pro-inflammatory Th1 to a more anti-inflammatory Th2 response (83). Unfortunately, therapy is associated with adverse reactions such as flu-like symptoms and transient laboratory abnormalities. Although IFNß therapy is widely used, the exact mechanism of action is poorly understood. Moreover, the response is only partial and a significant part of the patients fails to respond to treatment, the so-called non-responders (21). Other patients initially show a beneficial response to treatment, but after a prolonged period of treatment neutralizing antibodies (NAbs) may appear leading to a secondary non-response to treatment (84). Due to the lack of predictive biomarkers the question remains who will respond to IFNβ therapy and who to treat when inconvenience and costs are significant. In addition, other treatment strategies are needed aimed at inhibiting further disease progression and neurological damage.

Part of the unresponsiveness to IFNß can be explained by immunogenicity, i.e. the development of NAbs. However, since not all patients develop NAbs, and, if induced after at least six months of treatment, NAbs can disappear again over time (84-87), other mechanisms are likely to be involved to explain unresponsiveness. Hence, other mechanisms that result in insensitivity or resistance to the effects of IFNs are likely to underlie differential responsiveness. This implies that pharmacodynamic responses may differ between patients, leading to interindividual differences in clinical efficacy. Therefore, an in depth understanding of the pharmacodynamic factors underlying the therapeutic mechanisms and therapy unresponsiveness is the key for the identification of predictive markers.

1.5 Genomics in autoimmunity: Reducing complexity

There is growing evidence that both RA and MS are heterogeneous diseases. However, we generally refer to these diseases in terms of a group average, which may hamper progress of our understanding of the genetic basis, pathogenic mechanisms, and the treatment success for subsets of patients. Unfortunately, criteria for subclassification of patients e.g. to select those patients who will benefit from a specific treatment are currently lacking.

By definition, nearly every aspect of a disease phenotype should be represented in the pattern of genes and proteins that are expressed in the patient. In the last decade, the analysis of gene expression in tissues and cells has evolved from the analysis of a selected set of genes to an efficient high throughput whole-genome screening approach of potentially all genes expressed using microarray technology (Figure 2). By a global analysis of the genes that are expressed in cells and tissues of an individual under different conditions and during disease, we can build up "gene expression profiles (signatures)" which characterize the dynamic functioning of the genome under pathophysiological conditions.

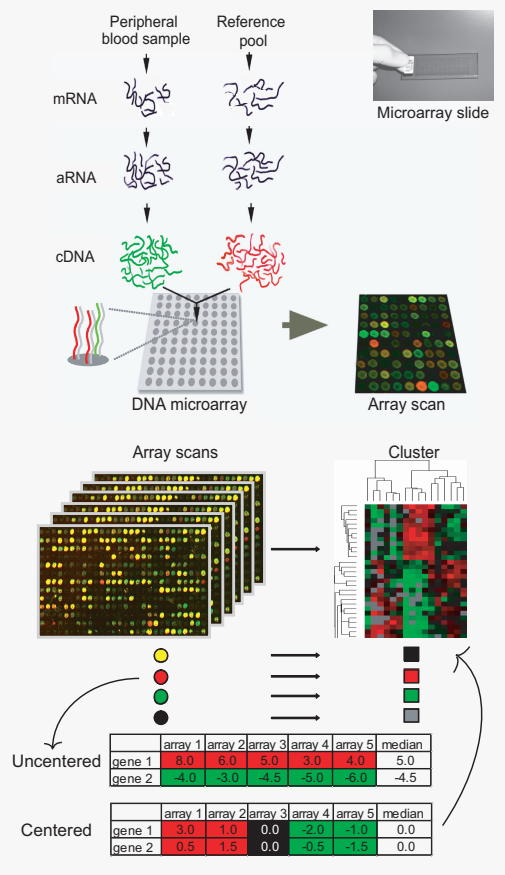


Figure 2. Microarray technology
RNA is isolated from an experimental sample and a control or reference sample and labelled with different fluorescent dyes. Both samples are then mixed together and hybridised onto the same microarray slide, resulting in a red to green ratio for each spotted element on the microarray. Scanning of the microarrays generates an image of all spots. The same common reference is used for all samples to allow comparison between the different arrays. Therefore, the data need to be centered i.e. ratio's are expressed relative to the median expression level across all arrays. For comparison of several samples a clustering algorithm can be used. Note the difference in colour representation of spots on the array and the corresponding squares in the cluster diagram.

1

1.5.1 Heterogenomics

The molecular signature typically represents the contributions and interactions of specific factors and distinct cells that are associated with disease characteristics and subtypes and thus defines the samples' unique biology. Identifying differential molecular signatures already in the preclinical phase of disease may provide a lead for the identification of predictive biomarkers for future disease development. This strategy also provides the means to subdivide patients that suffer from a complex heterogeneous disease into more homogeneous subgroups based on differences in gene expression profiles. The differentially expressed gene sets may then be used as prognostic markers for disease development, as disease classifiers, as surrogate marker for the measurement of a clinical endpoint and reflect the involvement of a particular biological pathway in disease (Figure 3). This discovery-based research identifies biological processes that may include new genes with unknown function or genes not previously known to be involved in this process. The latter category may hold surprises that sometimes urge us to redirect our thinking. Such a gene expression profiling approach has been proven useful in cancer research for the identification of classifiers for disease outcome and metastasis and underlying pathways (88;89). Hence, the identification of differentially expressed genes and proteins may provide a comprehensive molecular description of disease heterogeneity that will likely uncover biomarkers for future disease development and to allow stratification of patients for intervention therapies based on molecular criteria.

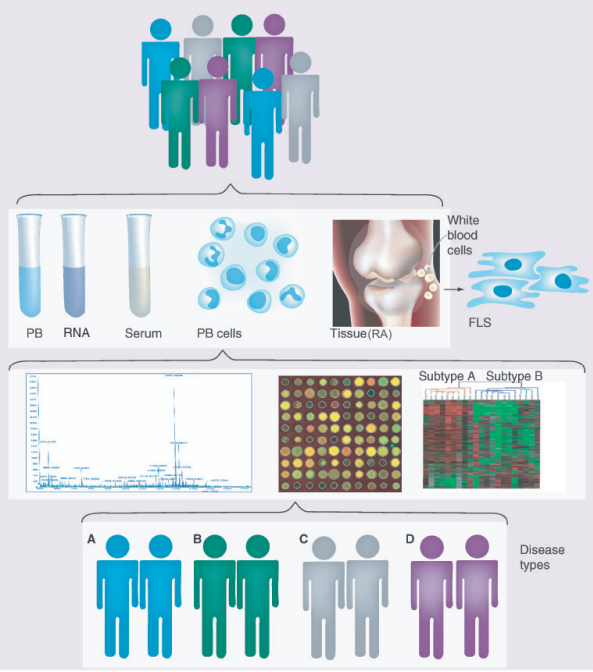


Figure 3. Heterogenomics
Application of technologies such as transcriptomics, proteomics and metabolomics allows us to determine the molecular description of an individual patient. When associated with clinical read-outs we could select the clinical useful molecular markers and apply these in day-to-day clinical practice.
The procedure starts with collecting from each patient different types of material like serum, peripheral blood (PB) cells, RNA from blood (using e.g. PAXgene tubes) and in the case of RA patients synovial biopsies and fibroblast-like synoviocytes (FLS) may be obtained. This material can then be analyzed by high-throughput techniques such as mass spectrometry, DNA-, protein- or antibody-arrays. Subsequently, computational

algorithms will be applied, to select subgroups of patients e.g. I. Purely based on differences in a molecular profile, II Based on the presence or absence of a known response signature for the target of interest or III. To identify markers significantly associated with a clinical phenotype or disease outcome. These data help to elucidate the distinct pathological mechanisms at play that can explain the inter-patient variation in future disease development, clinical presentation, disease progression, and treatment response. Knowledge of the differential pathogenic mechanisms helps us to identify new drug targets for selected patient subgroups. As an example, we hypothesize that a homogenous subgroup of patients could be stratified based on a specific molecular portrait that correlates with the responder/non-responder status of a targeted therapy. (Reproduced from Future Rheumatology June 2006 Vol. 1, No. 3, Pages 311-322, with permission of Future Medicine Ltd)

1.5.2 Pharmagenomics

Clinical experience in treating RA as well as MS patients indicates that there are responders and non-responders to treatment. The lack of clinical response to treatment suggests that some patients are insensitive or resistant to the action of treatment. This implies that pharmacodynamic responses may differ between patients, leading to interindividual differences in clinical response.

Receptor occupancy and subsequent initiation of a complex signal transduction program are the basis for the pharmacodynamic effects of cytokines (Figure 4). Ultimately, the cytokine-induced intracellular signalling program will switch on a wide variety of genes leading to specific transcriptional changes. The induction program of activated genes forms a rich source to identify genes that are suitable as pharmacodynamic marker of treatment, i.e. a marker that reflects the biological response to treatment. Thus, investigating changes in gene transcripts after adding or deleting a certain mediator of the immune response reflects the transcriptional response to that mediator.

The term pharmacogenomics emerged in the late 1990s and is associated with the application of genomics in drug development. Pharmacogenomics is defined as: "The investigation of variations of DNA and RNA characteristics as related to drug response". This applies to genetics and transcriptomics studies. With the aid of genomics technology, we are now in a position to provide sufficient knowledge to determine pharmacodynamic outcomes of biological treatments such as TNF antagonists and IFNβ. Longitudinal analysis of gene expression patterns from treated patients will provide an in-depth understanding of the pharmacodynamic consequences of biological therapy, and will allow us to evaluate the presence and/or absence of cytokine-induced pathways in relation to the clinical response to the drug.

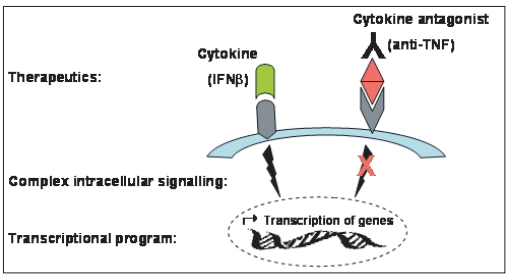


Figure 4.

Pharmacogenomic response to treatment

This scheme presents a simplistic overview of the effects of biologicals on receptor binding, signal transduction and gene program activation. The genes that are modulated upon treatment provide a framework to identify pharmacodynamic biomarkers to determine the drug effect on the target pathway.

1

1.6 Thesis outline

RA and MS patients reveal a striking heterogeneity based on clinical, biological and molecular criteria. Recognition of the preclinical phase of disease and categorization of patients is expected to be of utmost importance for decision making in clinical practice. Recent developments in high-throughput screening technologies have provided the opportunity to characterize patients based on their molecular profile. This thesis describes the application of genomics research in RA and MS for the identification of:

1. Biomarkers for preclinical diagnosis
2. Biomarkers for disease subclassification
3. New/unknown disease pathways that contribute to disease
4. Biomarkers for responsiveness to therapy and mechanisms of drug action

This information will improve our understanding of the underlying biology of the disease. Ultimately this information will help clinicians to select subgroups of patients for optimal treatment.

To reach these objectives MS and RA are used as two prototypically studied autoimmune diseases which may reflect two different ends of the autoimmune response. Studying both diseases may lead to a better understanding of the key players involved in autoimmunity in general and their specific roles in different autoimmune diseases. While affected synovial tissue from RA patients is relatively easy to obtain, the inability to collect brain or spinal cord biopsies from MS patients during disease, hinders genomics studies on CNS lesions in large patient cohorts. Based on the autoimmune nature of MS, and the migration of immune cells to and from the brain of affected individuals, it is hypothesized that processes that contribute to disease in the CNS are reflected in the periphery.

In chapter 2 we describe the use of genomics technology to improve the early diagnosis of disease. In chapter 3 we describe the use of genomics for subclassification of patients based on significant differences in gene expression profiles. Subsequently, in chapter 4 we investigate variations of RNA characteristics in relation to drug response, which is defined as the term “pharmacogenomics”. Finally, in chapter 5 the results presented in this thesis are summarized and discussed.

1.7 Reference List

(1) Hartley SB, Crosbie J, Brink R, Kantor AB, Basten A, Goodnow CC. Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. Nature 1991; 353(6346):765-769.

(2) William J, Euler C, Christensen S, Shlomchik MJ. Evolution of autoantibody responses via somatic hypermutation outside of germinal centers. Science 2002; 297(5589):2066-2070.

(3) Davidson A, Diamond B. Autoimmune diseases 1. N Engl J Med 2001; 345(5):340-350.

(4) Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive

review. Pharmacol Ther 2008; 117(2):244-279.

(5) Kollias G, Douni E, Kassiotis G, Kontoyiannis D. On the role of tumor necrosis factor and receptors in models of multiorgan failure, rheumatoid arthritis, multiple sclerosis and inflammatory bowel disease. Immunol Rev 1999; 169:175-194.

(6) Grell M, Wajant H, Zimmermann G, Scheurich P. The type 1 receptor (CD120a) is the high-affinity receptor for soluble tumor necrosis factor. Proc Natl Acad Sci U S A 1998; 95(2):570-575.

(7) Grell M, Douni E, Wajant H, Lohden M, Clauss M, Maxeiner B et al. The transmembrane form of tumor necrosis factor is the prime activating ligand of the 80 kDa tumor necrosis factor receptor. Cell 1995; 83(5):793-802.

(8) Pocsik E, Duda E, Wallach D. Phosphorylation of the 26 kDa TNF precursor in monocytic cells and in transfected HeLa cells. J Inflamm 1995; 45(3):152-160.

(9) Feldmann M, Maini RN. Lasker Clinical Medical Research Award. TNF defined as a therapeutic target for rheumatoid arthritis and other autoimmune diseases. Nat Med 2003; 9(10):1245-1250.

(10) ISAACS A, LINDENMANN J. Virus interference. I. The interferon. Proc R Soc Lond B Biol Sci 1957; 147(927):258-267.

(11) Pestka S, Krause CD, Walter MR. Interferons, interferon-like cytokines, and their receptors. Immunol Rev 2004; 202:8-32.

(12) Siegal FP, Kadowaki N, Shodell M, Fitzgerald-Bocarsly PA, Shah K, Ho S et al. The nature of the principal type 1 interferon-producing cells in human blood. Science 1999; 284(5421):1835-1837.

(13) Luft T, Pang KC, Thomas E, Hertzog P, Hart DN, Trapani J et al. Type I IFNs enhance the terminal differentiation of dendritic cells. J Immunol 1998; 161(4):1947-1953.

(14) Santini SM, Lapenta C, Logozzi M, Parlato S, Spada M, Di PT et al. Type I interferon as a powerful adjuvant for monocyte-derived dendritic cell development and activity in vitro and in Hu-PBL-SCID mice. J Exp Med 2000; 191(10):1777-1788.

(15) Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2003; 100(5):2610-2615.

(16) Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. Curr Opin Immunol 2004; 16(6):801-807.

(17) Braun D, Geraldes P, Demengeot J. Type I Interferon controls the onset and severity of autoimmune manifestations in lpr mice. J Autoimmun 2003; 20(1):15-25.

(18) Santiago-Raber ML, Baccala R, Haraldsson KM, Choubey D, Stewart TA, Kono DH et al. Type-I interferon receptor deficiency reduces lupus-like disease in NZB mice. J Exp Med 2003; 197(6):777-788.

(19) Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity 2006; 25(3):383-392.

(20) Nikpour M, Dempsey AA, Urowitz MB, Gladman DD, Barnes DA. Association of a gene expression profile from whole blood with disease activity in systemic lupus erythaematosus. Ann Rheum Dis 2008; 67(8):1069-1075.

(21) Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial.The IFNB Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. Neurology 1995; 45(7):1277-1285.

(22) van Holten J, Reedquist K, Sattonet-Roche P, Smeets TJ, Plater-Zyberk C, Vervoordeldonk MJ et al. Treatment with recombinant interferon-beta reduces inflammation and slows cartilage destruction in the collagen-induced arthritis model of rheumatoid arthritis. Arthritis Res Ther 2004; 6(3):R239-R249.

(23) van Holten J, Pavelka K, Vencovsky J, Stahl H, Rozman B, Genovese M et al. A multicentre, randomised, double blind, placebo controlled phase II study of subcutaneous interferon beta-1a in the treatment of patients with active rheumatoid arthritis. Ann Rheum Dis 2005; 64(1):64-69.

(24) van Oosten BW, Barkhof F, Truyen L, Boringa JB, Bertelsmann FW, von Blomberg BM et al. Increased MRI activity and immune activation in two multiple sclerosis patients treated with the monoclonal anti-tumor necrosis factor antibody cA2. Neurology 1996; 47(6):1531-1534.

(25) Sicotte NL, Voskuhl RR. Onset of multiple sclerosis associated with anti-TNF therapy. Neurology 2001; 57(10):1885-1888.

(26) Gomez-Reino JJ, Carmona L, Valverde VR, Mola EM, Montero MD. Treatment of rheumatoid arthritis with tumor necrosis factor inhibitors may predispose to significant increase in tuberculosis risk: a multicenter active-surveillance report. Arthritis Rheum 2003; 48(8):2122-2127.

(27) Ali Y, Shah S. Infliximab-induced systemic lupus erythematosus. Ann Intern Med 2002; 137(7):625-626.

(28) Bleumink GS, ter Borg EJ, Ramselaar CG, Ch SB. Etanercept-induced subacute cutaneous lupus erythematosus. Rheumatology (Oxford) 2001; 40(11):1317-1319.

(29) Mohan AK, Edwards ET, Cote TR, Siegel JN, Braun MM. Drug-induced systemic lupus erythematosus and TNF-alpha blockers. Lancet 2002; 360(9333):646.

(30) Shakoor N, Michalska M, Harris CA, Block JA. Drug-induced systemic lupus erythematosus associated with etanercept therapy. Lancet 2002; 359(9306):579-580.

(31) Charles PJ, Smeenk RJ, De Jong J, Feldmann M, Maini RN. Assessment of antibodies to double-stranded DNA induced in rheumatoid arthritis patients following treatment with infliximab, a monoclonal antibody to tumor necrosis factor alpha: findings in open-label and randomized placebo-controlled trials. Arthritis Rheum 2000; 43(11):2383-2390.

(32) Palucka AK, Blanck JP, Bennett L, Pascual V, Banchereau J. Cross-regulation of TNF and IFN-alpha in autoimmune diseases. Proc Natl Acad Sci U S A 2005; 102(9):3372-3377.

(33) Hofman FM, Hinton DR, Johnson K, Merrill JE. Tumor necrosis factor identified in multiple sclerosis brain. J Exp Med 1989; 170(2):607-612.

(34) Rothuizen LE, Buclin T, Spertini F, Trinchard I, Munafo A, Buchwalder PA et al. Influence of interferon beta-1a dose frequency on PBMC cytokine secretion and biological effect markers. J Neuroimmunol 1999; 99(1):131-141.

(35) Banchereau J, Pascual V, Palucka AK. Autoimmunity through cytokine-induced dendritic cell activation. Immunity 2004; 20(5):539-550.

(36) Lee DM, Weinblatt ME. Rheumatoid arthritis 1. Lancet 2001; 358(9285):903-911.

(37) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-324.

(38) Arnett FC. Revised criteria for the classification of rheumatoid arthritis. Bull Rheum Dis 1989; 38(5):1-6.

(39) Soden M, Rooney M, Cullen A, Whelan A, Feighery C, Bresnihan B. Immunohistological features in the synovium obtained from clinically uninvolved knee joints of patients with rheumatoid arthritis. Br J Rheumatol 1989; 28(4):287-292.

(40) Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000; 43(1):155-163.

(41) Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004; 50(2):380-386.

(42) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003; 48(10):2741-2749.

(43) Kuhn KA, Kulik L, Tomooka B, Braschler KJ, Arend WP, Robinson WH et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J Clin Invest 2006; 116(4):961-973.

(44) Lundberg K, Nijenhuis S, Vossenaar ER, Palmblad K, van Venrooij WJ, Klareskog L et al. Citrullinated proteins have increased immunogenicity and arthritogenicity and their presence in arthritic joints correlates with disease severity. Arthritis Res Ther 2005; 7(3):R458-R467.

(45) Quinn MA, Gough AK, Green MJ, Devlin J, Hensor EM, Greenstein A et al. Anti-CCP antibodies measured at disease onset help identify seronegative rheumatoid arthritis and predict radiological and functional outcome. Rheumatology (Oxford) 2006; 45(4):478-480.

(46) Wolfgang J.Hueber, Beren H.Tomooka, Kevin Deane, Lezlie A.Parrish, V.Michael Holers, William H.Robinson. Autoantibody Profiling in Pre-Disease RA Samples. Arthritis and Rheumatism . 2007. Ref Type: Abstract

(47) Rantapaa-Dahlqvist S, Boman K, Tarkowski A, Hallmans G. Up regulation of monocyte chemoattractant protein-1

expression in anti-citrulline antibody and immunoglobulin M rheumatoid factor positive subjects precedes onset of inflammatory response and development of overt rheumatoid arthritis. Ann Rheum Dis 2007; 66(1):121-123.

(48) Tak PP, Smeets TJ, Daha MR, Kluin PM, Meijers KA, Brand R et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. Arthritis Rheum 1997; 40(2):217-225.

(49) Klimiuk PA, Goronzy JJ, Bjor nJ, Beckenbaugh RD, Weyand CM. Tissue cytokine patterns distinguish variants of rheumatoid synovitis. Am J Pathol 1997; 151(5):1311-1319.

(50) Ulfgren AK, Grondal L, Lindblad S, Khademi M, Johnell O, Klareskog L et al. Interindividual and intra-articular variation of proinflammatory cytokines in patients with rheumatoid arthritis: potential implications for treatment. Ann Rheum Dis 2000; 59(6):439-447.

(51) Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM et al. Lymphoid neogenesis in rheumatoid synovitis. J Immunol 2001; 167(2):1072-1080.

(52) Berglin E, Padyukov L, Sundin U, Hallmans G, Stenlund H, van Venrooij WJ et al. A combination of autoantibodies to cyclic citrullinated peptide (CCP) and HLA-DRB1 locus antigens is strongly associated with future onset of rheumatoid arthritis. Arthritis Res Ther 2004; 6(4):R303-R308.

(53) van Gaalen FA, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL et al. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. Arthritis Rheum 2004; 50(3):709-715.

(54) Bongi SM, Manetti R, Melchiorre D, Turchini S, Boccaccini P, Vanni L et al. Anti-cyclic citrullinated peptide antibodies are highly associated with severe bone lesions in rheumatoid arthritis anti-CCP and bone damage in RA. Autoimmunity 2004; 37(6-7):495-501.

(55) van Venrooij WJ, Zendman AJ, Pruijn GJ. Autoantibodies to citrullinated antigens in (early) rheumatoid arthritis. Autoimmun Rev 2006; 6(1):37-41.

(56) Zendman AJW, Vossenaar ER, van Venrooij WJ. Autoantibodies to citrullinated (poly)peptides: A key diagnostic and prognostic marker for rheumatoid arthritis. Autoimmunity 2004; 37(4):295-299.

(57) Forslind K, Ahlmen M, Eberhardt K, Hafstrom I, Svensson B. Prediction of radiological outcome in early rheumatoid arthritis in clinical practice: role of antibodies to citrullinated peptides (anti-CCP). Ann Rheum Dis 2004; 63(9):1090-1095.

(58) Kastbom A, Strandberg G, Lindroos A, Skogh T. Anti-CCP antibody test predicts the disease course during 3 years in early rheumatoid arthritis (the Swedish TIRA project). Ann Rheum Dis 2004; 63(9):1085-1089.

(59) Ronnelid J, Wick MC, Lampa J, Lindblad S, Nordmark B, Klareskog L et al. Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-CP status predicts worse disease activity and greater radiological progression. Ann Rheum Dis 2005; 64(12):1744-1749.

(60) Raza K, Breese M, Nightingale P, Kumar K, Potter T, Carruthers DM et al. Predictive value of antibodies to cyclic citrullinated peptide in patients with very early inflammatory arthritis. J Rheumatol 2005; 32(2):231-238.

(61) Lindqvist E, Eberhardt K, Bendtzen K, Heinegard D, Saxne T. Prognostic laboratory markers of joint damage in rheumatoid arthritis. Ann Rheum Dis 2005; 64(2):196-201.

(62) De RL, Peene I, Hoffman IE, Kruithof E, Union A, Meheus L et al. Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: diagnostic value, associations with radiological progression rate, and extra-articular manifestations. Ann Rheum Dis 2004; 63(12):1587-1593.

(63) Elliott MJ, Maini RN, Feldmann M, Kalden JR, Antoni C, Smolen JS et al. Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor alpha (cA2) versus placebo in rheumatoid arthritis. Lancet 1994; 344(8930):1105-1110.

(64) Elliott MJ, Maini RN, Feldmann M, Long-Fox A, Charles P, Katsikis P et al. Treatment of rheumatoid arthritis with chimeric monoclonal antibodies to tumor necrosis factor alpha. Arthritis Rheum 1993; 36(12):1681-1690.

(65) Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. N Engl J Med 2003; 349(20):1907-1915.

(66) Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004; 350(25):2572-2581.

(67) McFarlin DE, McFarland HF. Multiple sclerosis (first of two parts). N Engl J Med 1982; 307(19):1183-1188.

(68) McFarlin DE, McFarland HF. Multiple sclerosis (second of two parts). N Engl J Med 1982; 307(20):1246-1251.

(69) McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol 2001; 50(1):121-127.

(70) Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple sclerosis. N Engl J Med 2000; 343(13):938-952.

(71) Hillert J, Gronning M, Nyland H, Link H, Olerup O. An immunogenetic heterogeneity in multiple sclerosis. J Neurol Neurosurg Psychiatry 1992; 55(10):887-890.

(72) Dyment DA, Ebers GC, Sadovnick AD. Genetics of multiple sclerosis. Lancet Neurol 2004; 3(2):104-110.

(73) Pugliatti M, Sotgiu S, Rosati G. The worldwide prevalence of multiple sclerosis. Clin Neurol Neurosurg 2002; 104(3):182-191.

(74) Hunter SF, Hafler DA. Ubiquitous pathogens: links between infection and autoimmunity in MS? Neurology 2000; 55(2):164-165.

(75) von Einsiedel RW, Samorei IW, Pawlita M, Zwissler B, Deubel M, Vinters HV. New JC virus infection patterns by in situ polymerase chain reaction in brains of acquired immunodeficiency syndrome patients with progressive multifocal leukoencephalopathy. J Neurovirol 2004; 10(1):1-11.

(76) Johnson RT. The virology of demyelinating diseases. Ann Neurol 1994; 36 Suppl:S54-S60.

(77) Soldan SS, Berti R, Salem N, Secchiero P, Flamand L, Calabresi PA et al. Association of human herpes virus 6 (HHV-6) with multiple sclerosis: increased IgM response to HHV-6 early antigen and detection of serum HHV-6 DNA. Nat Med 1997; 3(12):1394-1397.

(78) Wandinger K, Jabs W, Siekhaus A, Bubel S, Trillenberg P, Wagner H et al. Association between clinical disease activity and Epstein-Barr virus reactivation in MS. Neurology 2000; 55(2):178-184.

(79) Ferrante P, Mancuso R, Pagani E, Guerini FR, Calvo MG, Saresella M et al. Molecular evidences for a role of HSV-1 in multiple sclerosis clinical acute attack. J Neurovirol 2000; 6 Suppl 2:S109-S114.

(80) Croxford JL, Olson JK, Miller SD. Epitope spreading and molecular mimicry as triggers of autoimmunity in the Theiler's virus-induced demyelinating disease model of multiple sclerosis. Autoimmun Rev 2002; 1(5):251-260.

(81) Revesz T, Kidd D, Thompson AJ, Barnard RO, McDonald WI. A comparison of the pathology of primary and secondary progressive multiple sclerosis. Brain 1994; 117 (Pt 4):759-765.

(82) Thompson AJ, Polman CH, Miller DH, McDonald WI, Brochet B, Filippi MM, X et al. Primary progressive multiple sclerosis. Brain 1997; 120 (Pt 6):1085-1096.

(83) Yong VW, Chabot S, Stuve O, Williams G. Interferon beta in the treatment of multiple sclerosis: mechanisms of action. Neurology 1998; 51(3):682-689.

(84) Vartanian TK, Zamvil SS, Fox E, Sorensen PS. Neutralizing antibodies to disease-modifying agents in the treatment of multiple sclerosis. Neurology 2004; 63(11 Suppl 5):S42-S49.

(85) Sibley WA, Bamford CR, Clark K. Clinical viral infections and multiple sclerosis. Lancet 1985; 1(8441):1313-1315.

(86) Rice GP, Paszner B, Oger J, Lesaux J, Paty D, Ebers G. The evolution of neutralizing antibodies in multiple sclerosis patients treated with interferon beta-1b. Neurology 1999; 52(6):1277-1279.

(87) Reske D, Walser A, Haupt WF, Petereit HF. Long-term persisting interferon beta-1b neutralizing antibodies after discontinuation of treatment. Acta Neurol Scand 2004; 109(1):66-70.

(88) Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000; 403(6769):503-511.

(89) van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; %19;347(25):1999-2009.

Chapter 2
Genomics to improve diagnosis of disease

2

Chapter 2.1
Gene Expression Profiling in Autoantibody Positive Arthralgia Patients Predicts Development of Arthritis

Lisa G.M. van Baarsen[1], Wouter H. Bos, MD[2,3], François Rustenburg[1], Tineke C.T.M. van der Pouw Kraan, PhD[4], Gerrit Jan J. Wolbink, MD, PhD[2,3], Ben A.C. Dijkmans, MD, PhD[2,5], Dirkjan van Schaardenburg, MD, PhD[2,5], Cornelis L. Verweij, PhD[1,5]

[1]Dept. of Pathology, [4]Dept. of Molecular Cell Biology & Immunology, [5]Dept. of Rheumatology, VU University Medical Center, Amsterdam, Netherlands
[2]Jan van Breemen Institute, Amsterdam, Netherlands
[3]Sanquin Research, Amsterdam, Netherlands

Submitted for publication

Abstract

Objective: To identify molecular features that are associated with the development of rheumatoid arthritis (RA), to understand the pathophysiology of preclinical development and to assign predictive biomarkers.

Methods: We included 109 ACPA and/or RF positive arthralgia patients with absence of arthritis who were at risk for RA, and 25 RA patients. Gene expression profiles of blood samples were determined by DNA microarray analysis and qPCR. Significance Analysis of Microarrays and Prediction Analysis of Microarrays were used to determine significance and classifiers, respectively. Cox-regression hazard analysis was used to assess the risk for arthritis development.

Results: Gene expression profiling of blood cells revealed heterogeneity among 109 at risk arthralgia patients based on differential expression of immune related genes. A total of 20 at risk patients have developed arthritis after a median of 7 months. Here, we report for the first time gene signatures relevant to development of arthritis. Signatures significantly associated with arthritis development (hazard ratio 4.5; 95% C.I. 1.3-15.4; *P*=0.016) were involved in interferon (IFN) immunity, hematopoiesis and cell trafficking. These processes were reminiscent of those present in RA patients, implying that the preclinical phase of disease carries features of established disease. Genes involved in B-cell immunology were associated with protection from progression to arthritis.

Conclusion: Our results indicate that IFN-mediated immunity, hematopoiesis and cell trafficking specify processes relevant to progression to arthritis independent of ACPA positivity. These findings strongly suggest that certain gene signatures have predictive value for progression to arthritis, which will pave the way to preventive medicine.

Introduction

Rheumatoid arthritis (RA) is a destructive inflammatory autoimmune disease affecting mostly the joints. To induce remission and thereby prevent irreversible joint damage, early diagnosis and a timely start of effective treatment is of high importance (1-5).

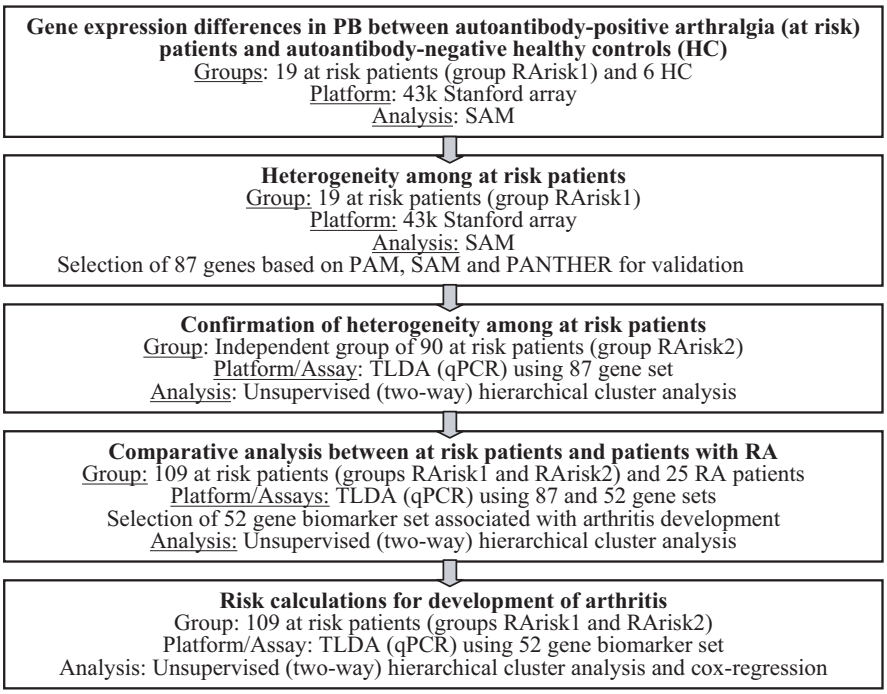
Alterations in the immune system are likely the basis for the development of RA. Accordingly, several studies have documented the appearance of antibodies against citrullinated proteins (ACPA) and rheumatoid factor (RF) prior to the onset of RA (6-11). Using serum samples stored in a blood bank, Rantapää-Dahlquist and colleagues showed that 34% of the RA patients were positive for ACPA up to nine years prior to diagnosis (9). In analogy, Nielen and colleagues showed that 49% of the RA patients tested positive for IgM-RF and/or ACPA before onset of disease

at a median of 4.5 years before symptom onset (10). A recent prospective follow-up study of ACPA and/or IgM RF positive arthralgia patients has shown that ACPA positive patients are more likely to develop arthritis than ACPA negative, IgM-RF positive arthralgia patients (27% versus 6% after a median follow-up of 2 years) (12). Above studies indicate that ACPA and/or RF may serve a role as predictive biomarkers for the development for RA in order to select individuals for preventive therapy for the development of RA.

It has been increasingly hypothesized that autoantibodies such as ACPA are involved in the development of RA (13-15). Animal studies demonstrated that addition of antibodies to citrullinated peptides significantly enhanced joint inflammation (14;16). Moreover, citrullination of collagen type II was shown to be more arthritogenic in rats (13). Since not all ACPA and/or RF positive individuals ultimately develop RA, the requirements to drive this process are likely to be different between the persons at risk (17;18). Hence, either additional factors are needed to result in a chronic inflammatory response ultimately leading to RA or some individuals may have a protective immune profile which suppresses disease development despite the presence of autoantibodies. The pathogenic or protective immune response might be selectively induced in susceptible individuals (19). Hueber et al. (20) suggested that anti-citrulline autoantibodies exhibit epitope spreading in the further preclinical and early phase of RA, indicative of affinity maturation. Additional research revealed an increase in the levels of monocyte chemoattractant protein-1 (MCP-1) in patients positive for ACPA before disease onset (21). These findings are supportive for a role of autoantibodies in combination with a specific immune constitution as early indicators for the development of RA. However, the exact nature of the pathogenic and/or protective response remains to be determined.

In the present explorative study we aim to identify molecular features in addition to autoantibodies that are associated with the development of (rheumatoid) arthritis in order to understand the pathobiological process and to identify additional biomarkers to improve the predictive power. Therefore, we applied peripheral blood gene expression profiling using a unique cohort of ACPA and/or IgM-RF positive arthralgia patients at risk for developing RA who are prospectively monitored for arthritis development. The flow of the different steps that we followed in this study is presented in figure 1.

Figure 1. Flowchart of data analysis



▪ Materials & Methods

Study population

Between June 2004 and March 2007, ACPA and/or IgM-RF positive arthralgia patients were included for prospective follow-up of arthritis development. Inclusion and exclusion criteria for this cohort have been described previously (22). In summary, a trained medical doctor (WB) and a senior rheumatologist (DS) independently scored for absence of arthritis (swollen joint count [SJC] =0) in 44 joints at physical examination at the baseline visit (23). The senior rheumatologist was blinded for the reported joint complaints and the autoantibody status. Exclusion criteria were: the presence of recent infections, autoimmune rheumatic diseases, cancer, arthritis revealed by chart review or baseline physical examination, erosions on hand or feet X-ray examination and previous treatment with a disease modifying anti-rheumatic drug (DMARD) or corticosteroids. In total, 109 patients were available for analysis. Arthritis development during follow-up was defined as a SJC of ≥ 1 and was independently confirmed by both physicians. The initial study was performed with peripheral blood samples of a consecutive series of 19 out of the 109 arthralgia patients and a group of 6 age, gender and ethnicity matched healthy controls who were ACPA and RF negative. For comparative studies, 25 randomly selected RA patients starting anti-TNF treatment at the Jan van Breemen Institute were

included. All fulfilled the ACR criteria for RA (24) and the median disease duration was 11 years. An overview of the patients' characteristics is given in Table 1.

Table 1. Clinical features of arthralgia and RA patients and healthy controls

	Arthralgia patients ACPA+ and/or RF+)		Healthy controls	RA patients*
Total (n)	19 (RA^{risk1})	109 ($RA^{risk1+risk2}$)	6	25
Age in years mean ± SD	48 ± 11	49 ± 10	53 ± 15	58 ± 12
Female (%)	13 (68)	75 (69)	4 (66)	19 (76)
ACPA positive, IgM-RF negative (%)	5 (26)	37 (34)	-	7 (28)
ACPA negative, IgM-RF positive (%)	9 (47)	39 (36)	-	0 (0)
ACPA and IgM-RF positive (%)	5 (26)	33 (30)	-	13 (52)

SD=standard deviation, ACPA=anti-citrullinated protein antibodies, IgM-RF = IgM rheumatoid factor, * Five RA patients were ACPA and RF negative

Serological measurements

Baseline laboratory parameters (determined batch wise at the end of the study period using the blood samples obtained at inclusion) included IgM-RF by in house enzyme-linked immunosorbent assay (ELISA) and ACPA by ELISA (second generation anti-CCP ELISA, Axis Shield, Dundee, United Kingdom). IgM-RF was calibrated with a national reference serum containing 200 IU/ml (25). The cut-off level for IgM-RF antibody positivity was set at 30 IU/ml determined on the basis of ROC curves as described previously (10). The cut-off level for ACPA positivity was set at 5 Arbitrary Units/ml (AU/ml) according to the manufacturer's instructions. High sensitive CRP levels were measured using CRPHS reagents with module C501 on a COBAS 6000 platform (Roche Diagnostics GmbH, Mannheim, Germany).

Blood sampling for RNA isolation

2.5 ml blood was drawn in PAXgene blood RNA isolation tubes (PreAnalytix, GmbH, Germany) and stored at -20°C. Tubes were thawed for 2 hours at room temperature prior to RNA isolation. Total RNA was isolated using the PAXgene RNA isolation kit according to the manufacturer's instructions including a DNAse (Qiagen, Venlo, Netherlands) step to remove genomic DNA.

Sample hybridization for microarray analysis

We used 43K cDNA microarrays from the Stanford Functional Genomics Facility (http://microarray.org/sfgf/) printed on aminosilane-coated slides containing ~20.000 unique genes. One batch of arrays was used for all experiments. First DNA spots were UV-crosslinked to the slide using 150-300 mJoules. Prior to sample hybridization, slides were pre-hybridized at 42 degrees Celsius for 15 minutes in a solution containing 40% ultra-pure formamide (Invitrogen, Breda, Netherlands), 5% SSC (Biochemika, Sigma), 0.1% SDS (Fluka Chemie, GmbH, Switzerland) and 50 g/ml BSA (Panvera, Madison, USA). After pre-hybridization slides were briefly rinsed in MilliQ water, thoroughly washed in boiling water and 95% ethanol and air-dried.

Sample preparation and microarray hybridization was performed as described previously (26) apart from the different post-processing and pre-hybridization described above.

Microarray data analysis

Data storage and filtering were performed using the Stanford Microarray Database (SMD at: http://genome-www5.stanford.edu//) (27) as described previously (28). Raw data can be downloaded from the publicly accessible Stanford database website. The Q-score tool from the SMD was used as a quality measure to remove low quality spots. Q-score determined the appropriate filter criteria for: the regression correlation between channels 1 and 2, the background settings and the minimal channel intensities. Data values with the same Unigene Identifier were averaged and all array data was median centered (genes and arrays) resulting in good quality data for 19,648 gene transcripts. Significance Analysis of Microarrays (SAM) (29) was used to determine significantly differential expressed genes. This algorithm corrects for multiple testing. A gene was considered as significantly differential expressed if the False Discovery Rate (FDR) was equal to or less than 5%. Cluster analysis (30) was used to study disease heterogeneity and to define clusters of coordinately expressed genes after which the data was visualized using Treeview. Prediction Analysis of Microarrays (PAM) was applied to assist in the selection of classifiers (31). PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (Applied Biosystems, Foster City, CA, USA) was used at http://PANTHER.appliedbiosystems.com (32;33) to interpret our data. This analysis uses the binomial statistics tool to compare the list of significantly up- or downregulated genes to a reference list in order to statistically determine over- or under representation of PANTHER classification categories such as biological processes. A Bonferroni correction was applied to correct for multiple testing and a significant p-value ($p<0.05$) indicates that a given category may be of biological interest.

Taqman® Low Density Arrays (TLDA)

Expression levels of selected target genes were validated in the whole study group (n=109) using TLDA (Applied Biosystems). Corresponding predesigned primers and probes (supplementary Table S1) were selected from the Applied Biosystems database to set up custom TLDA cards. Samples were randomly dispersed over the TLDA cards and the expression levels of selected genes were measured at the outsourcing company ServiceXS B.V. (Leiden, Netherlands). Total RNA (0.5 μg) was reverse transcribed into cDNA using a Revertaid H-minus cDNA synthesis kit (MBI Fermentas, St. Leon-Rot, Germany) according to the manufacturer's instructions. From each sample diluted cDNA corresponding to 100 ng total RNA was used per TLDA card. For normalization housekeeping genes GAPDH and 18SRNA were included in the analysis. Since the inter-sample variation in expression for GAPDH

was much higher than for 18SRNA, the latter was chosen for normalization. 65% of all genes quantified by TLDA showed a good correlation (Pearson R>0.4) with microarray data. An arbitrarily chosen ACPA-/RF- control sample was selected as calibrator sample and data was analyzed using RQ manager 1.2 (Applied Biosystems).

Statistical analysis

Data with a Gaussian distribution, expressed as the mean and SD, were analyzed using a T test or one-way ANOVA for multiple comparisons. Outcome measurements with a non-Gaussian distribution were expressed as the median and interquartile range (IQR) and were analyzed by the Mann-Whitney U test or Kruskal-Wallis test for multiple comparisons. Categorical variables were compared using the Chi-square or Fisher's exact test as appropriate. Cox-regression hazard analysis assessed the relative risk for arthritis development in subgroups of autoantibody positive arthralgia patients. To control for multiple comparisons, survival data were verified using a log rank analysis combined with a permutation approach. Data were analyzed using SPSS version 14.0 (Chicago, Illinois, United States) and were considered significant with two sided p-values less than 0.05.

Results

1. Gene expression profiles of ACPA and/or RF positive arthralgia patients are clearly distinct from autoantibody negative healthy controls

Gene expression profiles of peripheral blood cells derived from 19 arthralgia patients positive for ACPA and/or RF (group RA^{risk1}) were analyzed by microarray and compared to the profiles of 6 ACPA and RF negative healthy controls using SAM. A total of 255 genes were identified whose transcript levels showed significant differential expression of more than two-fold between the two groups (supplementary Table S2). Subsequently, unsupervised (two-way) hierarchical cluster analysis was applied, which revealed that all autoantibody negative healthy controls concentrated together in one arm of the dendogram (Fig. 2).

The autoantibody positive arthralgia patients were characterized by differential expression of two gene clusters designated as cluster A and cluster B. PANTHER classification analysis was used to interpret the biological function of the genes in these clusters. Cluster A consisted of genes that are upregulated in the risk group, and were involved in several immune related processes: interferon (IFN) mediated immunity, NK cell mediated immunity, and immunity and defense. Cluster B contained genes that are downregulated in most of the persons at risk. These genes could not be classified into a biological process.

Overall, peripheral blood gene expression profiles of autoantibody positive arthralgia patients were clearly distinct from those of autoantibody negative controls and revealed characteristics of an activated immune response.

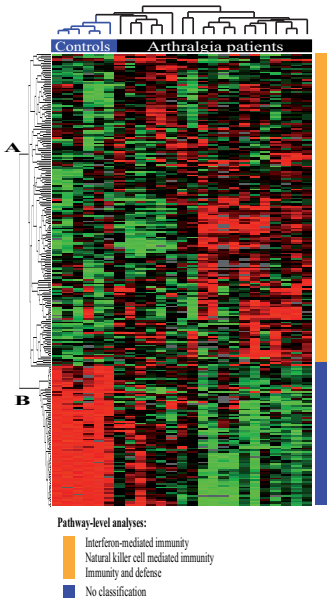


Figure 2. Cluster diagram of genes significantly differential expressed between autoantibody positive arthralgia patients and healthy controls
Supervised hierarchical cluster analysis of 255 genes that were significantly differential expressed between 19 autoantibody positive arthralgia patients and 6 healthy controls, identified by SAM (FDR <5%). The cluster diagram is visualized by Treeview. Each column represents the data of one array/sample and each row shows the relative expression level of a single gene for all samples. Red color means a relatively higher expression, green color stands for a relatively lower expression and black color indicates that the expression level is equal to the median expression level across all samples. Grey indicates missing values. The significantly differential expressed genes present in each gene cluster were classified into different functional categories (PANTHER) as indicated at the bottom of the cluster diagram.

2. Heterogeneity among autoantibody positive arthralgia patients
To investigate the molecular heterogeneity within the risk group, a total of 554 genes were identified whose transcript levels deviated more than two-fold from the median expression level in at least four out of the 19 at risk individuals. An unsupervised (two-way) hierarchical cluster analysis was performed based on the differential expression of the 554 gene set. The structure of the dendogram clearly indicated that the autoantibody positive arthralgia patients were separated in subgroups (Fig. 3).
The position of any autoantibody-positive arthralgia patient in the dendogram is determined by differentially expressed genes that are represented by four clusters (A, B, C and D). Pathway level analysis revealed that cluster A is characterized by the expression of genes involved in B-cell mediated immunity and protein biosynthesis, metabolism and modification. Genes involved in immunity and defense, proteolysis, protein metabolism and modification, macrophage, T-cell, NK-cell and granulocyte mediated immunity were characteristic for cluster B. Cluster C contained genes of yet unknown processes, although some of these genes are known for their role in inflammation (e.g. PBEF1, SSP1 and S100A12). Genes that represent immunity and defense, cell surface receptor and chemokine and cytokine mediated signaling, such as CCL5, IFNγ and IL32, were characteristic for cluster D.

3. Validation of heterogeneity among autoantibody positive arthralgia patients
The molecular heterogeneity was validated in an independent group of 90 autoantibody positive arthralgia patients (group RA^{risk2}). Therefore, we initially used a set of 87 genes that were representative for heterogeneity within the at risk

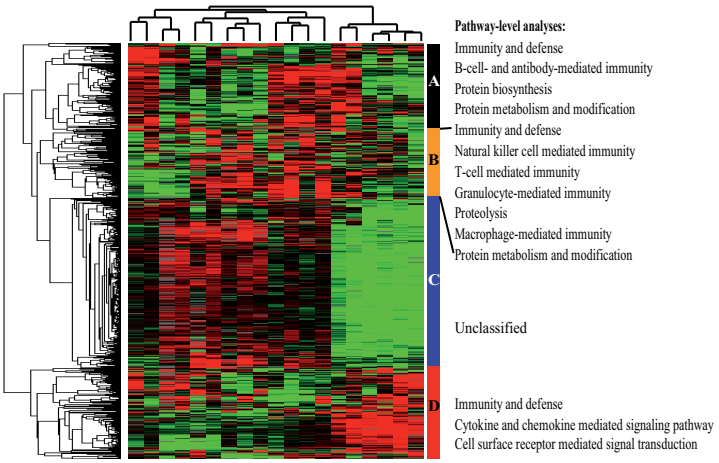


Figure 3. Molecular heterogeneity among autoantibody positive arthralgia patients
Two-way hierarchical cluster analysis of microarray data was used to investigate heterogeneity among peripheral blood gene expression profiles of 19 arthralgia patients. First, 554 genes were identified whose expression levels deviated from the median expression level at least twofold in at least four patients. Subsequently, these genes were unsupervised hierarchically clustered resulting in a subclassification of patients based on the differential expression of genes. The cluster diagram is visualized by Treeview. Each column represents the data of one array/sample and each row shows the relative expression level of a single gene for all samples. Red color means a relatively higher expression, green color stands for a relatively lower expression and black color indicates that the expression level is equal to the median expression level across all samples. Grey indicates missing values. The genes that are differentially expressed between patient subgroups fall in clusters A, B, C and D and were classified into different functional categories (PANTHER) as indicated at the bottom of the cluster diagram. The gene clusters A, B, C and D are color-coded and shown on the side.

group based on microarray data. These genes were selected by PAM and SAM and supplemented with genes representing the different identified biological processes. The gene set was supplemented with a previously described RA-associated set of 20 type I IFN response genes (26) and *IFNα2, IFNβ and IFNγ*. Detailed information on the criteria for gene selection is described in the supplementary information and Tables S3 and S4. The 87 genes were measured by qPCR (TLDA) and subsequent unsupervised (two-way) hierarchical cluster analysis confirmed the existence of heterogeneity among autoantibody positive arthralgia patients (supplementary Fig. S1).

In aggregate, molecular heterogeneity among autoantibody positive arthralgia patients based on differential expression of genes that are involved in diverse arms of the immune response is confirmed in an independent group of 90 autoantibody positive patients at risk.

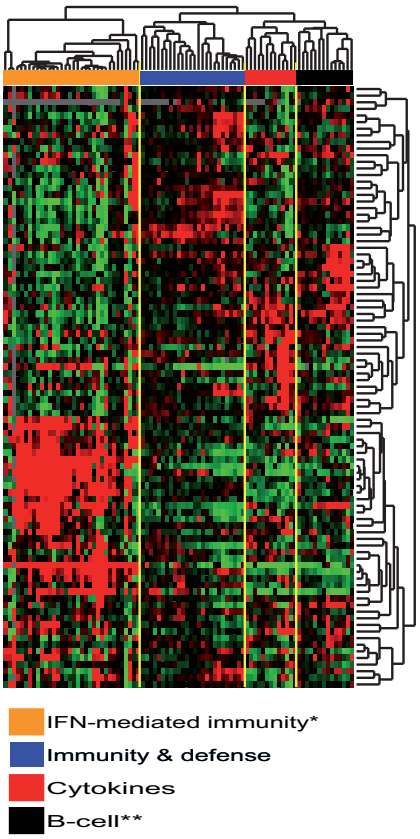


Figure S1. Confirmation of heterogeneity in an independent group of autoantibody positive arthralgia patients using TLDA

Unsupervised hierarchical cluster analysis of 87 genes measured by TLDA (qPCR) in an independent group of 90 autoantibody positive arthralgia patients (group RA^{risk2}). The cluster diagram is visualized by Treeview. Each row represents the expression data for one arthralgia patient and each column shows the relative expression level of a single gene for all patients. Genes and patients with similar expression profiles are placed next to each another. For each gene cluster pathway-level analysis (PANTHER) was used to determine the presence of biological processes that were significantly represented.

* This gene cluster also consists of genes involved in macrophage-mediated immunity and apoptosis.

** This gene cluster also consists of genes involved in signal transduction.

4. Comparative analysis between blood gene expression profiles of autoantibody positive arthralgia patients and patients with RA

Next, the existence of commonalities between (subsets of) autoantibody positive arthralgia patients and patients with established RA was investigated using gene expression values obtained by TLDA. Therefore, we performed a comparative analysis between the gene expression characteristics of whole blood from all the 109 at risk individuals (group RA^{risk1} and RA^{risk2}) and those of patients with established RA (n=25). In order to reduce the gene set we selected those 52 genes (Table S5) out of the 87 genes that showed significantly differential expression levels between the patient subgroups.

Unsupervised hierarchical cluster analysis using the 52 gene set divided the arthralgia and RA patients into four subgroups (Fig. 4A) based on essentially the same biological processes as mentioned earlier i.e. IFN-mediated immunity, B-cell activation, cytokine/chemokine mediated immunity, apoptosis and NK cell mediated immunity. In addition, a small cluster of genes represented the biological process hematopoiesis. Most remarkable, patients with established RA preferentially

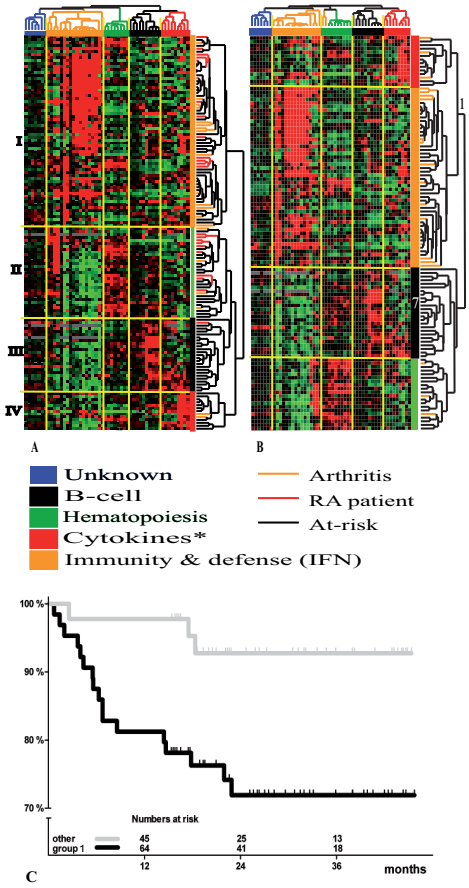


Figure 4. Gene expression profiles associated with arthritis development

A. Unsupervised hierarchical cluster analysis of gene expression data of the 52 gene set measured by TLDA, of autoantibody positive arthralgia patients who have not (black labels) and who have developed arthritis (orange labels), and patients with established RA (red labels). Each column represents the relative expression level of a single gene for all samples and each row shows the data of one patient's sample. Red color means a relatively higher expression, green color stands for a relatively lower expression and black color indicates that the expression level is equal to the median expression level across all samples. Grey indicates missing values. Genes are organized in color-coded clusters (see bottom of the diagram). The patients are categorized in groups I, II, III and IV (left side) as indicated by the dendograms and color codes on the right side. The biological themes of the gene clusters are indicated at the bottom of the cluster diagram. RA patients preferentially co-clustered with arthralgia patients present in subgroups I and II.

B. Unsupervised hierarchical cluster analysis of the 52 gene set performed for the at risk arthralgia patients only. Subgroups 1 and 7 associated with arthritis development are indicated on the right side.

C. Kaplan-Meyer curve of arthritis-free survival in subgroup 1 of the at risk arthralgia patients as determined in figure 4B. Cox-regression analysis revealed that subgroup 1 is significantly associated with arthritis development (HR 4.5; 95% C.I. 1.3-15.4; P=0.016).

co-clustered with at risk individuals of groups I and II compared to those in groups III and IV (Fisher's exact test $P<0.01$). Groups I and II were characterized by an increased expression of genes involved IFN-mediated immunity and hematopoiesis, respectively, whereas groups III and IV were associated with increased expression of genes involved in B cell activation and cytokine/chemokine mediated immunity, respectively.

These data show that a subgroup of autoantibody positive arthralgia patients contains peripheral blood gene expression features, i.e. increased expression of genes involved IFN-mediated immunity and/or hematopoiesis that are reminiscent of those of patients with established RA.

5. Gene expression profiles predict arthritis development, independent of ACPA levels

In order to study the association between the arthralgia subgroups and the development of arthritis we performed an analysis to determine the distribution

over the different subgroups of those autoantibody positive arthralgia patients who have developed arthritis in the course of the study. Interim analysis revealed that 20 autoantibody positive arthralgia patients have developed arthritis after a median of 7 months (IQR 4-15; median follow-up of all patients is 30 [IQR 22-39] months) in a median of 3 joints (IQR 3-5).

Since autoantibody positive arthralgia patients present in subgroups I and II co-clustered with RA patients, these two subgroups were combined and arthritis development was compared to the remaining at risk patients present in subgroups III and IV. Cox-regression analysis showed that subgroups I and II were associated with arthritis development (Hazard Ratio [HR] 5.1; 95% confidence interval [C.I.] 1.2-21.9; P=0.03). Correcting for ACPA decreased the HR to 4.1 (95% C.I. 1.0-17.9; P=0.06), resulting in a strong trend independent of ACPA status, implicating that in the presence of ACPA, gene expression profiles specific for subgroups I and II have predictive value for identifying those patients at risk for the development of RA. The presence of the 'shared epitope' genotype, as well as mean ACPA, RF and CRP levels were similar in both arthralgia subgroups (I and II versus III and IV, data not shown) and did not influence these results. Excluding those autoantibody positive arthralgia patients who had received two intramuscular dexamethasone injections (n=25) in a trial of primary prevention of arthritis (34) did not alter these results.

Essentially similar results were found when the two-way hierarchical cluster analysis was exclusively performed with the at risk arthralgia patients (Fig. 4B). However, since after re-clustering of the autoantibody positive arthralgia patients there is no objective criterion to decide which patient groups should be tested for association with arthritis development, multiple cuts in the dendogram were considered. Therefore, all clusters containing more than 15% of the 109 patients were tested versus the other clusters combined for association with arthritis development. This resulted in nine two-group tests. Cox-regression analysis revealed that subgroup 1, characterized by an increased expression of genes involved in IFN-mediated immunity and/or cytokine/chemokine mediated immunity, is significantly associated with arthritis development (Fig. 4C; HR 4.5; 95% C.I. 1.3-15.4; *P*=0.016) even after correcting for ACPA (HR 3.6; 95% C.I. 1.1-12.4; *P*=0.041). Subgroup 7, characterized by increased expression of genes involved in B-cell mediated immunity, is devoid of at risk individuals who developed arthritis. Therefore, inclusion of a "dummy" case was required for Cox regression analysis, which resulted in borderline significance (HR 0.14; 95% C.I. 0.02-1.1; *P*=0.057), which was lost after correcting for ACPA (HR 0.17; 95% C.I. 0.02-1.3; *P*=0.083). A log rank analysis combined with a permutation approach to control for multiple comparisons showed similar results (data not shown).

Collectively, these analyses reveal that autoantibody positive arthralgia patients with high an expression of genes involved in IFN-mediated immunity, cytokine/chemokine mediated immunity or hematopoiesis are more likely to develop arthritis.

Conversely, autoantibody positive arthralgia patients with high expression of genes involved in B-cell mediated immunity may be protected against development of arthritis or progression towards disease pathogenesis may be suppressed. These results were independent of ACPA status.

Discussion

In this manuscript we demonstrated the heterogeneous nature of ACPA and/or RF positive arthralgia patients at risk for development of RA. We identified sets of genes whose expression profiles segregate arthralgia patients at risk for RA into different subgroups. Table 2 summarizes the relative risk of arthritis development in subgroups of autoantibody positive arthralgia patients. Subgroups that are characterized by a gene signature of IFN-mediated immunity, cytokine activity, or hematopoiesis all contain at risk persons who have developed arthritis. These gene expression characteristics increase the risk for arthritis development approximately 4-fold independent of ACPA status. Interestingly, the group of patients that is characterized by increased expression of genes involved in humoral immunity is devoid of patients who have developed arthritis in the follow-up period. These results indicate that predisposition for the development of arthritis can be used to predict the diagnosis of arthritis in ACPA and/or RF positive individuals at risk.

Table 2. Relative risk of arthritis development in subgroups of autoantibody positive arthralgia patients

Figure	At risk subgroups	Arthritis development n (% of total risk subgroup)	HR	HR after ACPA correction	Conclusion
4A	I+II vs. III+IV	18 (23%) vs. 2 (5%)	5.1*	4.1	Arthralgia patients clustering together with RA patients have an increased risk of developing arthritis
4B	1. vs. rest	17 (27%) vs. 3 (7%)	4.5*	3.6*	Arthralgia patients with an increased expression of genes involved in IFN-mediated immunity, cytokine mediated immunity or hematopoiesis have an increased risk of developing arthritis
4B	7 vs. rest	0 (0%) vs. 20 (24%)	0.14^	0.17^	Arthralgia patients with an increased expression of genes involved in B-cell activation are protected from developing arthritis.

* P-value < 0.05
^ A "dummy" case that developed arthritis was required for Cox regression analysis
Subgroups: I. IFN-mediated immunity, II. Hematopoiesis, III. B-cell activity, IV. Cytokines
Subgroups: 1. Cytokines and IFN-mediated immunity, 7. B-cell activity

To the best of our knowledge, this study is the first to present evidence that differential gene expression is predictive of the progression to arthritis in

autoantibody positive individuals. Since the unique nature of this cohort prevents validation in an independent group, this predictor gene set is only suggestive and awaits confirmation from larger independent data sets to refine and validate its clinical usefulness. In addition, it would be interesting to re-analyze the results after a longer follow-up period.

The concept that IFN response activity in blood cells is characteristic for progression to arthritis is consistent with its role in immunopathology. Type I IFNs are known to upregulate MHC expression and to induce differentiation of monocytes into antigen presenting dendritic cells (DCs) (35;36). Under normal physiology immature DCs control peripheral tolerance by deletion of circulating autoreactive T cells. However, continued type I IFN induced maturation of DCs may lead to a break of peripheral tolerance through activation of autoreactive cells resulting in immunity to self-antigens. Accordingly, a recent study showed that IFN induced protein IFIT4 might play a role in promoting monocyte differentiation into DC-like cells and subsequent regulation of Th1 cell differentiation (37). These properties of type I IFN might contribute to antigen spreading and subsequent initiation of arthritis in those autoantibody positive arthralgia patients who display increased levels of IFN induced genes.

Among the risk factors for progression to arthritis are also genes involved in cytokine and chemokine mediated immunity, including IFNγ, IL-7R, STAT-4 and CCL5. This finding is consistent with the results of a study on undifferentiated arthritis, which showed correlations between raised concentrations of multiple cytokines and the presence of ACPA in early RA (38). In addition Hueber and colleagues observed raised serum levels of both classical Th1 (IFNγ and IL12) and Th2 (IL10 and IL13) cytokines (39). Whereas serum cytokine levels may reflect local blood cell production and/or (immune) cell activation in joints or other tissues, the increased transcript production in blood is indicative of peripheral production of these inflammatory factors. The presence of increased STAT-4 may be a reflection of IL-12 bioactivity and Th1 skewing reflected by increased expression of IFNγ. STAT-4 is also required for signaling in mature dendritic cells in response to type 1 IFNs (40). In accordance with a role for STAT-4 in progression towards RA, a genetic association between a STAT-4 haplotype and RA susceptibility was reported (41). Although previous findings suggested that increased levels of pro-inflammatory cytokines and/or chemokines are associated with the generation of ACPA (39), our findings suggest that increased cytokine activity is associated with progression towards arthritis in already ACPA positive individuals.

The group of genes classified as being involved in hematopoiesis contains two genes (KLF1 and ERAF) involved in erythrocyte development, whereas two other genes (BCL2L1 and BAG1) have anti-apoptotic capacities. Interestingly, this cluster also harbors DARC (Duffy Antigen/Receptor for Chemokines), also known as Duffy blood group antigen which is entirely unique among all the chemokine

receptors and is expressed on erythrocytes of Duffy antigen positive individuals. On erythrocytes, DARC is described to act as a sink for free pro-inflammatory chemokines present in the bloodstream (42). In addition, erythrocyte DARC may also serve as a chemokine blood reservoir that could have consequences for cell trafficking (43).
Compared with at risk individuals who do progress to arthritis, the ones that do not develop arthritis showed an increased expression of B-cell related genes. Notably transcripts for CD79A, CD79B, MS4A1, CD19 and FCLR5 were differentially expressed. CD79A and CD79B are B-cell markers essential for the expression and function of the B-cell antigen receptor. MS4A1 encodes for the B-cell antigen CD20, and CD19 is involved in activating the B-cell receptor. FCRL5 inhibits B-cell activation and is found on most mature B-cells with the highest levels being present on naïve and plasma cells. The finding that a gene set representing increased humoral immunity has a protective role against progression to arthritis is in apparent contradiction to the profound disease-ameliorating effects that have been observed following B-cell depletion (44). As a consequence of increased trafficking and accumulation of activated and autoreactive B-cells in the synovial membrane of affected joints, diminished numbers of circulating CD19 B-cells have been observed in RA patients (45). Knowing that immune cell trafficking is a crucial process in the initiation of disease, it is tempting to speculate that the increased expression of B-cell genes is a reflection of a block in the migration of B-cells to the tissue.

On the basis of our data, we propose that the individual's genes affect susceptibility to RA at three levels. First, some genes affect the overall reactivity of the immune system and thus can predispose the individual to autoimmunity. Second, this altered immunoreactivity is directed to particular antigens, i.e. citrullinated antigens, which affect B- and T-cell recognition of epitopes. Third, still other genes act on the progression of autoimmunity to target tissues to modulate immune attack. Our results imply that, among others IFN-mediated immunity and cell trafficking specify the processes relevant to progression to arthritis besides autoantibody positivity.

Can gene expression profiling be used to select those individuals at risk who will develop arthritis? 27% of ACPA positive arthralgia patients develop arthritis after a median follow-up of 2 years, this increases to 40% with the concomitant presence of IgM-RF (46). Our result suggests that higher-order combinatorial searches may improve the predictive performance of autoantibody status. However, higher-order predictive variable combinations do require the support of many more samples to prevent overfitting of the model. Convincing assessment of this question will therefore necessitate expression analysis of genes identified herein in larger cohorts of participants. We envision predictive models based on preclinical expression profiling as an "evolving" evidence-based process for determining

the risk of developing RA, to be recalibrated over time to account for changes in practice. As a corollary, a gene set that can identify RA progressors could also have predictive value for autoantibody negative patients. Studies are under way to further characterize all the genes in the classifier and to refine our predictor model with a larger patient cohort, with the ultimate goal of diagnosing patients with preclinical RA. Preclinical diagnosis could represent the basis of a breakthrough in curing RA.

Funding:

This study was supported by the Netherlands Organization for Health Research and Development (ZonMW grant number 6100.0010), the European Community's FP6 funding (AutoCure) and the Centre for Medical Systems Biology (Netherlands Genomics Initiative).

Acknowledgements

The authors would like to thank Drs Jose Ferreira and Mark van de Wiel from the department of Clinical Epidemiology & Biostatistics of the VUMC, Amsterdam for their expert advice in statistics.

Reference List

(1) Quinn MA, Conaghan PG, Emery P. The therapeutic approach of early intervention for rheumatoid arthritis: what is the evidence? Rheumatology (Oxford) 2001; 40(11):1211-1220.

(2) O'Dell JR. Treating rheumatoid arthritis early: a window of opportunity? Arthritis Rheum 2002; 46(2):283-285.

(3) Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. Arthritis Rheum 2002; 46(4):894-898.

(4) Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, van ZD, Kerstens PJ, Hazes JM et al. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. Ann Intern Med 2007; 146(6):406-415.

(5) Emery P, McInnes IB, van VR, Kraan MC. Clinical identification and treatment of a rapidly progressing disease state in patients with rheumatoid arthritis. Rheumatology (Oxford) 2008; 47(4):392-398.

(6) Aho K, Palosuo T, Raunio V, Puska P, Aromaa A, Salonen JT. When does rheumatoid disease start? Arthritis Rheum 1985; 28(5):485-489.

(7) Aho K, von ER, Kurki P, Palosuo T, Heliovaara M. Antikeratin antibody and antiperinuclear factor as markers for subclinical rheumatoid disease process. J Rheumatol 1993; 20(8):1278-1281.

(8) Aho K, Palosuo T, Heliovaara M, Knekt P, Alha P, von ER. Antifilaggrin antibodies within "normal" range predict rheumatoid arthritis in a linear fashion. J Rheumatol 2000; 27(12):2743-2746.

(9) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003; 48(10):2741-2749.

(10) Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis

Rheum 2004; 50(2):380-386.

(11) Jorgensen KT, Wiik A, Pedersen M, Hedegaard CJ, Vestergaard BF, Gislefoss RE et al. Cytokines, autoantibodies and viral antibodies in premorbid and postdiagnostic sera from patients with rheumatoid arthritis: case-control study nested in a cohort of Norwegian blood donors. Ann Rheum Dis 2008; 67(6):860-866.

(12) Bos WH, Wolbink GJ, Boers M, Tijhuis GJ, de VN, van der Horst-Bruinsma IE et al. Arthritis development in arthralgia patients is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. Ann Rheum Dis 2009.

(13) Lundberg K, Nijenhuis S, Vossenaar ER, Palmblad K, van Venrooij WJ, Klareskog L et al. Citrullinated proteins have increased immunogenicity and arthritogenicity and their presence in arthritic joints correlates with disease severity. Arthritis Res Ther 2005; 7(3):R458-R467.

(14) Kuhn KA, Kulik L, Tomooka B, Braschler KJ, Arend WP, Robinson WH et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J Clin Invest 2006; 116(4):961-973.

(15) Quinn MA, Gough AK, Green MJ, Devlin J, Hensor EM, Greenstein A et al. Anti-CCP antibodies measured at disease onset help identify seronegative rheumatoid arthritis and predict radiological and functional outcome. Rheumatology (Oxford) 2006; 45(4):478-480.

(16) Hill JA, Bell DA, Brintnell W, Yue D, Wehrli B, Jevnikar AM et al. Arthritis induced by posttranslationally modified (citrullinated) fibrinogen in DR4-IE transgenic mice. J Exp Med 2008; 205(4):967-979.

(17) Klareskog L, Alfredsson L, Rantapaa-Dahlqvist S, Berglin E, Stolt P, Padyukov L. What precedes development of rheumatoid arthritis? Ann Rheum Dis 2004; 63 Suppl 2:ii28-ii31.

(18) Klareskog L, Ronnelid J, Lundberg K, Padyukov L, Alfredsson L. Immunity to citrullinated proteins in rheumatoid arthritis. Annu Rev Immunol 2008; 26:651-675.

(19) Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. J Immunol 2003; 171(2):538-541.

(20) Wolfgang J.Hueber, Beren H.Tomooka, Kevin Deane, Lezlie A.Parrish, V.Michael Holers, William H.Robinson. Autoantibody Profiling in Pre-Disease RA Samples. Arthritis and Rheumatism . 2007. Ref Type: Abstract

(21) Rantapaa-Dahlqvist S, Boman K, Tarkowski A, Hallmans G. Up regulation of monocyte chemoattractant protein-1 expression in anti-citrulline antibody and immunoglobulin M rheumatoid factor positive subjects precedes onset of inflammatory response and development of overt rheumatoid arthritis. Ann Rheum Dis 2007; 66(1):121-123.

(22) Bos WH, Ursum J, de VN, Bartelds GM, Wolbink GJ, Nurmohamed MT et al. The role of the shared epitope in arthralgia with anti-cyclic citrullinated peptide antibodies (anti-CCP), and its effect on anti-CCP levels. Ann Rheum Dis 2008; 67(9):1347-1350.

(23) van der Heijde DM, van 't HM, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol 1993; 20(3):579-581.

(24) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-324.

(25) Klein F, Janssens MB. Standardisation of serological tests for rheumatoid factor measurement. Ann Rheum Dis 1987; 46(9):674-680.

(26) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis 2007; 66(8):1008-1014.

(27) Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D et al. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. Nucleic Acids Res 2007; 35(Database issue): D766-D770.

(28) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006; 7(6):522-531.

(29) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-5121.

(30) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-14868.

(31) Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002; 99(10):6567-6572.

(32) Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13(9):2129-2141.

(33) Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res 2006; 34(Web Server issue):W645-W650.

(34) Bos WH, Dijkmans BA, Boers M, van de SR, van SD. Effect of dexamethasone on autoantibody levels and arthritis development in arthralgia patients: a randomized trial. Ann Rheum Dis 2009.

(35) Luft T, Pang KC, Thomas E, Hertzog P, Hart DN, Trapani J et al. Type I IFNs enhance the terminal differentiation of dendritic cells. **J Immunol 1998; 161(4):1947-1953.**

(36) Santini SM, Lapenta C, Logozzi M, Parlato S, Spada M, Di PT et al. **Type I interferon as a powerful adjuvant for monocyte-**derived dendritic cell development and activity in vitro and in Hu-PBL-SCID mice. J Exp Med 2000; 191(10):1777-1788.

(37) Huang X, Shen N, Bao C, Gu Y, Wu L, Chen S. Interferon-induced protein IFIT4 is associated with systemic lupus erythematosus and promotes differentiation of monocytes into DC-like cells. Arthritis Res Ther 2008; 10(4):R91.

(38) Stabler T, Piette JC, Chevalier X, Marini-Portugal A, Kraus VB. Serum cytokine profiles in relapsing polychondritis suggest monocyte/macrophage activation. Arthritis Rheum 2004; 50(11):3663-3667.

(39) Hueber W, Tomooka BH, Zhao X, Kidd BA, Drijfhout JW, Fries JF et al. Proteomic analysis of secreted proteins in early rheumatoid arthritis: anti-citrulline autoreactivity is associated with up regulation of proinflammatory cytokines. Ann Rheum Dis 2007; 66(6):712-719.

(40) Fukao T, Frucht DM, Yap G, Gadina M, O'Shea JJ, Koyasu S. Inducible expression of Stat4 in dendritic cells and macrophages and its critical role in innate and adaptive immune responses. J Immunol 2001; 166(7):4446-4455.

(41) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007; 357(10):977-986.

(42) Pruenster M, Rot A. Throwing light on DARC. Biochem Soc Trans 2006; 34(Pt 6):1005-1008.

(43) Fukuma N, Akimitsu N, Hamamoto H, Kusuhara H, Sugiyama Y, Sekimizu K. A role of the Duffy antigen for the maintenance of plasma chemokine concentrations. Biochem Biophys Res Commun 2003; 303(1):137-139.

(44) Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004; 350(25):2572-2581.

(45) Wagner U, Kaltenhauser S, Pierer M, Wilke B, Arnold S, Hantzschel H. B lymphocytopenia in rheumatoid arthritis is associated with the DRB1 shared epitope and increased acute phase response. Arthritis Res 2002; 4(4):R1.

(46) Bos WH, Boers M, Wolbink G, Tijhuis G, de Vries N, Bruinsma IV et al. Arthritis development in prospectively followed arthralgia patients is dependent on anti-citrullinated protein antibody status. Arthritis Rheum 2008; 58(9):S290.

Supplementary information

Table S1. TLDA assays

Symbol	AssayID	Gene Name
18SRNA	Hs99999901_s1	ribosomal 18S-specific
BAG1	Hs00185390_m1	BCL2-associated athanogene
BCL2L1	Hs00169141_m1	BCL2-like 1
C12orf35	Hs00216848_m1	chromosome 12 open reading frame 35
C18orf17	Hs00400521_m1	chromosome 18 open reading frame 17
CCL5	Hs00174575_m1	chemokine (C-C motif) ligand 5
CCT4	Hs00272345_m1	chaperonin containing TCP1, subunit 4 (delta)
CD19	Hs00174333_m1	CD19 molecule
CD274	Hs00204257_m1	CD274 molecule
CD79A	Hs00998119_m1	CD79a molecule
CD79B	Hs01058826_g1	CD79b molecule
CKS1B	Hs01029137_g1	CDC28 protein kinase regulatory subunit 1B
DARC	Hs01011079_s1	Duffy blood group
DDX5	Hs00189323_m1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
DEFA3	Hs00414018_m1	defensin alpha 3
DTX3L	Hs00370540_m1	deltex 3-like (Drosophila)
EEF1G	Hs01922638_u1	eukaryotic translation elongation factor 1 gamma
EIF2AK2 =PKR	Hs00169345_m1	eukaryotic translation initiation factor 2-alpha kinase 2
EPSTI1	Hs00264424_m1	epithelial stromal interaction 1 (breast)
ERAF	Hs00372339_g1	erythroid associated factor
FCGR1A	Hs00417598_m1	Fc fragment of IgG, high affinity Ia, receptor (CD64)
FCRL5	Hs00258709_m1	Fc receptor-like 5
FLJ31033	Hs00291459_m1	hypothetical protein FLJ31033
FLT3LG	Hs00181740_m1	fms-related tyrosine kinase 3 ligand
GADD45B	Hs00169587_m1	growth arrest and DNA-damage-inducible beta
GAPDH	Hs00266705_g1	glyceraldehyde-3-phosphate dehydrogenase
GBP1	Hs00266717_m1	guanylate binding protein 1, interferon-inducible, 67kDa
GZMH	Hs00277212_m1	granzyme H
HDAC10	Hs00368899_m1	histone deacetylase 10
HERC3	Hs00205934_m1	hect domain and RLD 3
HLA-G	Hs00365950_g1	histocompatibility antigen, class I, G
ID1	Hs00357821_g1	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein
IFI27	Hs00271467_m1	interferon, alpha-inducible protein 27
IFI44L	Hs00199115_m1	interferon-induced protein 44-like
IFI6	Hs00242571_m1	interferon, alpha-inducible protein 6
IFIH1 =MDA5	Hs00223420_m1	interferon induced with helicase C domain 1
IFIT1	Hs01911452_s1	interferon-induced protein with tetratricopeptide repeats 1
IFIT2	Hs00533665_m1	interferon-induced protein with tetratricopeptide repeats 2
IFITM1	Hs00705137_s1	interferon induced transmembrane protein 1 (9-27)

Symbol	AssayID	Gene Name
IFNA2	Hs00265051_s1	interferon, alpha 2
IFNAR2	Hs01022059_m1	interferon (alpha beta and omega) receptor 2
IFNB1	Hs01077958_s1	interferon, beta 1, fibroblast
IFNG	Hs00174143_m1	interferon, gamma
IGKC	Hs00415165_m1	immunoglobulin kappa constant, immunoglobulin kappa variable 1-5
IGLL1	Hs00252263_m1	immunoglobulin lambda-like polypeptide 1
IL1A	Hs00174092_m1	interleukin 1 alpha
IL32	Hs00170403_m1	interleukin 32
IL7R	Hs00902334_m1	interleukin 7 receptor
IPO7	Hs00255188_m1	importin 7
IRF2	Hs00180006_m1	interferon regulatory factor 2
ISG15	Hs00192713_m1	ISG15 ubiquitin-like modifier
KDELR3	Hs00423556_m1	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
KLF1	Hs00610592_m1	Kruppel-like factor 1 (erythroid)
LGALS3BP	Hs00174774_m1	lectin, galactoside-binding, soluble, 3 binding protein
LTB	Hs00242739_m1	lymphotoxin beta (TNF superfamily, member 3)
LTF	Hs00914330_m1	lactotransferrin
MAT2B	Hs00203231_m1	methionine adenosyltransferase II, beta
MMP9	Hs00234579_m1	matrix metallopeptidase 9 (gelatinase B, 92kDa)
MRPL38	Hs00375656_m1	mitochondrial ribosomal protein L38
MS4A1	Hs00544819_m1	membrane-spanning 4-domains, subfamily A, member 1
MX1	Hs00182073_m1	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)
MYOM2	Hs00187676_m1	myomesin (M-protein) 2, 165kDa
NKG7	Hs00366585_g1	natural killer cell group 7 sequence
OAS1	Hs00242943_m1	oligoadenylate synthetase 1, 40/46 kDa
OAS2	Hs00159719_m1	oligoadenylate synthetase 2, 69/71 kDa
OAS3	Hs00196324_m1	oligoadenylate synthetase 3, 100 kDa
PADI2	Hs00247108_m1	peptidyl arginine deiminase, type II
PARP14	Hs00393814_m1	poly (ADP-ribose) polymerase family, member 14
PBEF1	Hs00237184_m1	pre-B-cell colony enhancing factor 1
PDK3	Hs00178440_m1	pyruvate dehydrogenase kinase isozyme 3
PLSCR1	Hs00275514_m1	phospholipid scramblase 1
RBM38	Hs00250139_m1	RNA binding motif protein 38
REN	Hs00166915_m1	renin
RGS18	Hs00329468_m1	regulator of G-protein signaling 18
RPL23	Hs00745462_s1	ribosomal protein L23
RSAD2	Hs00369813_m1	radical S-adenosyl methionine domain containing 2
RUFY1	Hs00228528_m1	RUN and FYVE domain containing 1
S100A12	Hs00194525_m1	S100 calcium binding protein A12
S100A8	Hs00374264_g1	S100 calcium binding protein A8
SAMD9L	Hs00541567_s1	sterile alpha motif domain containing 9-like

Symbol	AssayID	Gene Name
SELENBP1	Hs00187625_m1	selenium binding protein 1
SELL	Hs00174151_m1	selectin L (lymphocyte adhesion molecule 1)
SERPING1	Hs00163781_m1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)
SLC25A1	Hs00761590_sH	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1
SPP1	Hs00959010_m1	secreted phosphoprotein 1 (osteopontin)
STAT1	Hs00234829_m1	signal transducer and activator of transcription 1, 91kDa
STAT4	Hs00231372_m1	signal transducer and activator of transcription 4
TESC	Hs00215487_m1	tescalcin
TNFSF10 =TRAIL	Hs00234356_m1	tumor necrosis factor (ligand) superfamily, member 10
TNFSF7	Hs00174297_m1	CD70 molecule
TPM3	Hs00383595_m1	tropomyosin 3
TRGV9	Hs00233330_m1	TCR gamma alternate reading frame protein, T cell receptor gamma variable 9
TRIM22	Hs00232319_m1	tripartite motif-containing 22
XRCC5	Hs00221707_m1	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen

Table S2. Significantly differential expressed genes between arthralgia patients and controls

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.17518 || RSAD2 || Radical S-adenosyl methionine domain containing 2	Hs.17518	5.04	0.00
|| || || AI347124 || || || IMAGE:1926927 || 10229	IMAGE:1926927	4.84	0.00
Hs.389724 || IFI44L || Interferon-induced protein 44-like	Hs.389724	4.72	0.14
Hs.546467 || EPSTI1 || Epithelial stromal interaction 1 (breast)	Hs.546467	4.49	0.07
Hs.562188 || || Transcribed locus	Hs.562188	4.45	0.00
Hs.105635 || || Transcribed locus	Hs.105635	4.25	0.14
Hs.20315 || IFIT1 || Interferon-induced protein with tetratricopeptide repeats 1	Hs.20315	3.73	0.00
Hs.591198 || EVI2A || Ecotropic viral integration site 2A	Hs.591198	3.59	0.00
Hs.528634 || OAS3 || 2'-5'-oligoadenylate synthetase 3, 100kDa	Hs.528634	3.41	0.00
Hs.567266 || NF1 || Neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)	Hs.567266	3.36	0.07
Hs.534346 || RPS7 || Ribosomal protein S7	Hs.534346	3.29	0.00
Hs.567702 || || **Transcribed locus, strongly similar to NP_000997.1 ribosomal protein S3a; 40S ribosomal protein S3a; v-fos transformation effector protein 1 [Homo sapiens]	Hs.567702	3.26	0.09
Hs.23956 || COMMD8 || COMM domain containing 8	Hs.23956	3.13	0.00
Hs.610444 || RPL9 || Ribosomal protein L9	Hs.610444	3.13	0.11
Hs.406300 || RPL23 || Ribosomal protein L23	Hs.406300	3.06	0.00
Hs.226307 || APOBEC3B || Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B	Hs.226307	3.05	0.40
Hs.448851 || USP6 || **Ubiquitin specific peptidase 6 (Tre-2 oncogene)	Hs.448851	3.03	0.00

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.541894 || KIAA1641 || KIAA1641	Hs.541894	3.02	0.00
|| || || AA259248::AA259249 || || || IMAGE:1010072 || 17970	IMAGE:1010072	3.01	0.07
Hs.279806 || DDX5 || DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	Hs.279806	2.98	0.00
Hs.507536 || MTMR6 || **Myotubularin related protein 6	Hs.507536	2.95	0.00
Hs.643209 || || Transcribed locus	Hs.643209	2.93	0.18
Hs.531067 || || CDNA clone IMAGE:4902949	Hs.531067	2.88	0.05
Hs.445129 || C12orf35 || Chromosome 12 open reading frame 35	Hs.445129	2.88	0.00
Hs.530436 || STXBP3 || Syntaxin binding protein 3	Hs.530436	2.84	0.00
Hs.384598 || SERPING1 || Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	Hs.384598	2.78	1.21
Hs.510402 || CD46 || CD46 molecule, complement regulatory protein	Hs.510402	2.69	0.00
Hs.163173 || IFIH1 || Interferon induced with helicase C domain 1	Hs.163173	2.67	0.00
Hs.307924 || MSL3L1 || Male-specific lethal 3-like 1 (Drosophila)	Hs.307924	2.66	0.00
Hs.301921 || CCR1 || Chemokine (C-C motif) receptor 1	Hs.301921	2.66	0.00
|| || || AA912032 || || || IMAGE:1486028 || 101762	IMAGE:1486028	2.66	0.00
Hs.642906 || || Transcribed locus	Hs.642906	2.65	0.00
Hs.24485 || SMC3 || Chondroitin sulfate proteoglycan 6 (bamacan)	Hs.24485	2.64	0.00
|| || || AI268082 || || || IMAGE:1911561 || 100393	IMAGE:1911561	2.63	0.00
Hs.25845 || MGC42105 || Hypothetical protein MGC42105	Hs.25845	2.63	1.51
|| || || AI302425 || || || IMAGE:1901759 || 5293	IMAGE:1901759	2.61	0.00
Hs.470943 || STAT1 || Signal transducer and activator of transcription 1, 91kDa	Hs.470943	2.61	0.00
|| || || AA778116 || || || IMAGE:379941 || 19701	IMAGE:379941	2.61	0.00
Hs.34576 || TAX1BP1 || Tax1 (human T-cell leukemia virus type I) binding protein 1	Hs.34576	2.60	0.00
Hs.489118 || SAMD9L || Sterile alpha motif domain containing 9-like	Hs.489118	2.59	0.00
Hs.597910 || || Transcribed locus (1)	Hs.597910	2.58	0.00
Hs.643431 || IGJ || Immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides)	Hs.643431	2.56	0.11
Hs.381123 || RPL21 || **Ribosomal protein L21	Hs.381123	2.56	0.00
Hs.546523 || || Full-length cDNA clone CS0DK002YF13 of HeLa cells Cot 25-normalized of Homo sapiens (human)	Hs.546523	2.55	0.00
Hs.35804 || HERC3 || Hect domain and RLD 3	Hs.35804	2.55	0.00
Hs.430849 || OSBPL8 || Oxysterol binding protein-like 8	Hs.430849	2.53	0.00
Hs.437855 || FLJ21908 || Hypothetical protein FLJ21908	Hs.437855	2.52	0.00
|| || || AA502230 || || || IMAGE:911901 || 1872	IMAGE:911901	2.49	0.79
Hs.440534 || TXNDC10 || Thioredoxin domain containing 10	Hs.440534	2.47	0.00
Hs.622965 || MOBKL1A || **MOB1, Mps One Binder kinase activator-like 1A (yeast)	Hs.622965	2.47	0.21
Hs.276770 || CD52 || CD52 molecule	Hs.276770	2.45	0.69

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.643515 || BIRC2 || Baculoviral IAP repeat-containing 2	Hs.643515	2.43	0.00
Hs.468840 || PLEK || Pleckstrin	Hs.468840	2.41	0.52
Hs.306769 || RUFY1 || RUN and FYVE domain containing 1	Hs.306769	2.41	0.00
Hs.501778 || TRIM22 || Tripartite motif-containing 22	Hs.501778	2.40	0.00
Hs.277937 || GZMK || Granzyme K (granzyme 3; tryptase II)	Hs.277937	2.40	0.11
Hs.524760 || OAS1 || 2',5'-oligoadenylate synthetase 1, 40/46kDa	Hs.524760	2.38	0.00
Hs.177861 || SF3B14 || Splicing factor 3B, 14 kDa subunit	Hs.177861	2.37	0.05
Hs.250009 || ARL8B || ADP-ribosylation factor-like 8B	Hs.250009	2.36	0.00
Hs.387787 || KLRK1 || Killer cell lectin-like receptor subfamily K, member 1	Hs.387787	2.35	0.00
Hs.432996 || FLJ11021 || Similar to splicing factor, arginine/serine-rich 4	Hs.432996	2.35	0.00
Hs.476052 || SNRK || SNF related kinase	Hs.476052	2.35	0.00
Hs.82848 || SELL || Selectin L (lymphocyte adhesion molecule 1)	Hs.82848	2.35	0.00
Hs.593163 || || CDNA FLJ42968 fis, clone BRSTN2016954	Hs.593163	2.35	0.00
Hs.138701 || TRAT1 || T cell receptor associated transmembrane adaptor 1	Hs.138701	2.35	0.00
Hs.643522 || IPO7 || Importin 7	Hs.643522	2.33	0.00
Hs.205163 || MRPL3 || Mitochondrial ribosomal protein L3	Hs.205163	2.33	0.00
Hs.601000 || || Transcribed locus	Hs.601000	2.33	0.00
Hs.339453 || ERGIC2 || ERGIC and golgi 2	Hs.339453	2.32	0.00
Hs.295626 || ITGB1 || Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	Hs.295626	2.32	0.00
Hs.355809 || LOC151507 || Similar to male-specific lethal 3-like 1 isoform a; drosophila MSL3-like 1	Hs.355809	2.32	0.00
Hs.599179 || || CDNA FLJ37302 fis, clone BRAMY2016009	Hs.599179	2.31	0.00
Hs.529439 || ZBTB41 || Zinc finger and BTB domain containing 41	Hs.529439	2.31	0.18
Hs.634689 || || Transcribed locus	Hs.634689	2.29	0.00
Hs.81848 || RAD21 || RAD21 homolog (S. pombe)	Hs.81848	2.29	0.00
Hs.7886 || PELI1 || Pellino homolog 1 (Drosophila)	Hs.7886	2.29	0.00
Hs.518201 || DTX3L || Deltex 3-like (Drosophila)	Hs.518201	2.29	0.00
Hs.505326 || NELL2 || NEL-like 2 (chicken)	Hs.505326	2.28	0.00
Hs.523332 || OAT || Ornithine aminotransferase (gyrate atrophy)	Hs.523332	2.28	0.05
|| || || AA971523::AI733335::AI791673 || || || IMAGE:1584391 || 115069	IMAGE:1584391	2.28	0.00
Hs.85769 || DNTTIP2 || Deoxynucleotidyltransferase, terminal, interacting protein 2	Hs.85769	2.27	0.00
|| || || AI061169::AI733618 || || || IMAGE:1699682 || 56377	IMAGE:1699682	2.27	0.00
Hs.547382 || || CDNA clone IMAGE:30389268	Hs.547382	2.25	0.00
Hs.478150 || PDCD10 || Programmed cell death 10	Hs.478150	2.24	0.07
Hs.494173 || ANXA1 || Annexin A1	Hs.494173	2.24	0.34
Hs.512743 || || CDNA FLJ37755 fis, clone BRHIP2023762	Hs.512743	2.24	0.00
Hs.3210 || REN || Renin	Hs.3210	2.24	0.00
Hs.364544 || TSPAN13 || Tetraspanin 13	Hs.364544	2.24	0.52

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.591157 || KLRC2 || Killer cell lectin-like receptor subfamily C, member 2	Hs.591157	2.24	0.21
Hs.596900 || || Full-length cDNA clone CS0DB005YH06 of Neuroblastoma Cot 10-normalized of Homo sapiens (human)	Hs.596900	2.23	0.28
Hs.530472 || || Transcribed locus, weakly similar to NP_055301.1 neuronal thread protein AD7c-NTP [Homo sapiens]	Hs.530472	2.23	0.00
Hs.236774 || HMGN4 || High mobility group nucleosomal binding domain 4	Hs.236774	2.22	0.00
Hs.130413 || TM9SF2 || Transmembrane 9 superfamily member 2	Hs.130413	2.22	0.00
Hs.438801 || || Transcribed locus, strongly similar to NP_001735.1 calcium/calmodulin-dependent protein kinase IV; brain Ca(2+)-calmodulin-dependent protein kinase type IV; calcium/calmodulin-dependent protein kinase	Hs.438801	2.21	0.00
Hs.546339 || C11orf58 || **Chromosome 11 open reading frame 58	Hs.546339	2.21	0.00
Hs.370487 || B4GALT5 || UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 5	Hs.370487	2.21	0.00
Hs.535011 || FLJ31033 || Hypothetical protein FLJ31033	Hs.535011	2.20	0.00
Hs.191518 || DHX9 || DEAH (Asp-Glu-Ala-His) box polypeptide 9	Hs.191518	2.19	0.00
Hs.309090 || SFRS7 || Splicing factor, arginine/serine-rich 7, 35kDa	Hs.309090	2.18	0.00
Hs.477495 || MYSM1 || Myb-like, SWIRM and MPN domains 1	Hs.477495	2.18	0.00
Hs.528836 || NOD27 ||Nucleotide-binding oligomerization domain 27	Hs.528836	2.17	0.00
Hs.504609 || ID1 || Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein	Hs.504609	2.17	0.00
Hs.593155 || UTRN || Utrophin (homologous to dystrophin) (2)	Hs.593155	2.17	0.00
Hs.77424 || FCGR1A || Fc fragment of IgG, high affinity Ia, receptor (CD64)	Hs.77424	2.17	0.69
Hs.435967 || TADA1L || Transcriptional adaptor 1 (HFI1 homolog, yeast)-like	Hs.435967	2.16	0.00
Hs.591232 || || Full-length cDNA clone CS0DE014YC17 of Placenta of Homo sapiens (human)	Hs.591232	2.16	0.00
Hs.518475 || EIF4A2 || Eukaryotic translation initiation factor 4A, isoform 2	Hs.518475	2.16	0.00
Hs.632478 || LOC653805 || Similar to ribosomal protein S27	Hs.632478	2.16	0.52
Hs.495960 || ATP6AP2 || **ATPase, H+ transporting, lysosomal accessory protein 2	Hs.495960	2.15	0.00
Hs.632532 || HAT1 || Histone acetyltransferase 1	Hs.632532	2.15	0.00
|| || || N54914 || || || IMAGE:244637 || 56837	IMAGE:244637	2.15	0.00
Hs.546283 || RPL6 || Ribosomal protein L6	Hs.546283	2.15	0.00
Hs.62661 || GBP1 || Guanylate binding protein 1, interferon-inducible, 67kDa	Hs.62661	2.14	0.00
Hs.528638 || ATPBD1C || ATP binding domain 1 family, member C	Hs.528638	2.14	0.00
Hs.210385 || HERC1 || **Hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RCC1 (CHC1)-like domain (RLD) 1	Hs.210385	2.13	0.00
Hs.508769 || PCID2 || PCI domain containing 2	Hs.508769	2.12	0.00
Hs.593056 || || Transcribed locus	Hs.593056	2.12	0.40
Hs.281898 || AIM2 || Absent in melanoma 2	Hs.281898	2.11	0.00

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.412103 || EFHA1 || EF-hand domain family, member A1	Hs.412103	2.11	0.00
Hs.597455 || || Transcribed locus	Hs.597455	2.11	0.05
Hs.518609 || ARPC5 || Actin related protein 2/3 complex, subunit 5 16kDa	Hs.518609	2.11	0.34
Hs.600495 || || Transcribed locus	Hs.600495	2.11	0.11
Hs.546510 || SLFN5 || **Schlafen family member 5	Hs.546510	2.11	0.05
Hs.86984 || BCOR || BCL6 co-repressor	Hs.86984	2.11	0.00
Hs.515890 || YPEL5 || Yippee-like 5 (Drosophila)	Hs.515890	2.11	0.28
Hs.608192 || || Transcribed locus	Hs.608192	2.11	0.11
Hs.123464 || P2RY5 || Purinergic receptor P2Y, G-protein coupled 5	Hs.123464	2.11	0.11
Hs.31961 || || CDNA FLJ37694 fis, clone BRHIP2015224	Hs.31961	2.11	0.00
Hs.79136 || SLC39A6 || Solute carrier family 39 (zinc transporter), member 6	Hs.79136	2.10	0.00
Hs.1908 || PRG1 || Proteoglycan 1, secretory granule	Hs.1908	2.09	0.07
Hs.508010 || FNDC3A || Fibronectin type III domain containing 3A	Hs.508010	2.09	0.00
Hs.530275 || || Transcribed locus, moderately similar to NP_060846.2 nipsnap homolog 3B [Homo sapiens]	Hs.530275	2.09	0.00
Hs.531664 || || Transcribed locus, strongly similar to XP_290670.5 PREDICTED: KIAA0220-like protein [Homo sapiens]	Hs.531664	2.09	0.00
Hs.632235 || SHBG || Sex hormone-binding globulin	Hs.632235	2.08	0.00
Hs.532634 || IFI27 || Interferon, alpha-inducible protein 27	Hs.532634	2.08	2.55
Hs.485380 || DBF4 || DBF4 homolog (S. cerevisiae)	Hs.485380	2.08	0.09
Hs.505874 || TBK1 || TANK-binding kinase 1	Hs.505874	2.07	0.00
Hs.467701 || ODC1 || Ornithine decarboxylase 1	Hs.467701	2.07	0.00
Hs.403010 || || **Transcribed locus, strongly similar to NP_057698.2 TRAF and TNF receptor-associated protein; ETS1-associated protein 2 [Homo sapiens]	Hs.403010	2.07	0.00
Hs.369284 || C20orf6 || Chromosome 20 open reading frame 6	Hs.369284	2.06	0.00
Hs.634736 || || Transcribed locus	Hs.634736	2.06	0.00
Hs.183684 || EIF4G2 || Eukaryotic translation initiation factor 4 gamma, 2	Hs.183684	2.06	0.00
Hs.606459 || || **Transcribed locus	Hs.606459	2.06	0.00
Hs.501140 || KIAA1598 || KIAA1598	Hs.501140	2.06	1.00
Hs.46578 || DOCK10 || Dedicator of cytokinesis 10	Hs.46578	2.06	0.00
Hs.432424 || TPP2 || Tripeptidyl peptidase II	Hs.432424	2.06	0.00
Hs.362807 || || MRNA; cDNA DKFZp667P0610 (from clone DKFZp667P0610)	Hs.362807	2.05	0.00
|| || || AA780560 || || || IMAGE:867580 || 95697	IMAGE:867580	2.05	0.00
Hs.106688 || CXorf1 || **Chromosome X open reading frame 1	Hs.106688	2.05	0.00
Hs.54642 || MAT2B || Methionine adenosyltransferase II, beta	Hs.54642	2.05	0.00
Hs.583792 || PARP14 || Poly (ADP-ribose) polymerase family, member 14	Hs.583792	2.05	0.00
Hs.584801 || SFRS2 || Splicing factor, arginine/serine-rich 2	Hs.584801	2.05	0.00
Hs.156171 || PSMC6 || Proteasome (prosome, macropain) 26S subunit, ATPase, 6	Hs.156171	2.05	0.00

Genes higher expressed in arthralgia patients (181):

Gene Name	Gene ID	Fold Change	q-value (%)
Hs.508148 || ABI1 || Abl-interactor 1	Hs.508148	2.04	0.00
Hs.600488 || || Transcribed locus	Hs.600488	2.04	0.11
Hs.311346 || CMAS || Cytidine monophosphate N-acetylneuraminic acid synthetase	Hs.311346	2.04	0.00
Hs.558348 || ITK || IL2-inducible T-cell kinase	Hs.558348	2.04	0.00
Hs.534255 || B2M || Beta-2-microglobulin	Hs.534255	2.04	0.00
Hs.512576 || KLRC1 || Killer cell lectin-like receptor subfamily C, member 1	Hs.512576	2.04	0.34
Hs.465323 || KIAA1468 || KIAA1468	Hs.465323	2.03	0.05
Hs.633281 || || Transcribed locus	Hs.633281	2.03	0.18
Hs.156625 || C1orf80 || Chromosome 1 open reading frame 80	Hs.156625	2.03	0.00
Hs.601789 || || Transcribed locus	Hs.601789	2.03	0.00
Hs.513440 || GPR65 || G protein-coupled receptor 65	Hs.513440	2.02	0.05
Hs.594861 || || CDNA FLJ13601 fis, clone PLACE1010069	Hs.594861	2.02	0.00
Hs.118041 || UBP1 || Upstream binding protein 1 (LBP-1a)	Hs.118041	2.02	0.00
Hs.159028 || BTN2A1 || Butyrophilin, subfamily 2, member A1	Hs.159028	2.02	0.00
Hs.445239 || HOXA2 || Homeobox A2	Hs.445239	2.02	0.00
Hs.138378 || CASP4 || Caspase 4, apoptosis-related cysteine peptidase	Hs.138378	2.02	0.00
Hs.202010 || PLCL2 || Phospholipase C-like 2	Hs.202010	2.01	0.00
Hs.512152 || HLA-G || HLA-G histocompatibility antigen, class I, G	Hs.512152	2.01	4.88
Hs.33922 || C1orf156 || Chromosome 1 open reading frame 156	Hs.33922	2.01	0.11
Hs.558798 || TRIM73 || Tripartite motif-containing 73	Hs.558798	2.01	0.79
Hs.522863 || CYorf15A || Chromosome Y open reading frame 15A	Hs.522863	2.01	0.18
Hs.471040 || FLJ38973 || Hypothetical protein FLJ38973	Hs.471040	2.00	0.00
Hs.631513 || REST || RE1-silencing transcription factor	Hs.631513	2.00	0.00

Genes lower expressed in arthralgia patients (79)

Gene name	Gene ID	Fold Change	q-value (%)
Hs.125713 || FAM92B || Family with sequence similarity 92, member B	Hs.125713	0.13	0.79
Hs.116210 || LOC199899 || Hypothetical protein LOC199899	Hs.116210	0.17	2.55
Hs.404997 || TRIM62 || Tripartite motif-containing 62	Hs.404997	0.25	1.00
Hs.435462 || DEADC1 || Deaminase domain containing 1	Hs.435462	0.25	4.21
Hs.443683 || MYOM2 || Myomesin (M-protein) 2, 165kDa	Hs.443683	0.26	1.51
Hs.515550 || LOC284361 || Hematopoietic signal peptide-containing	Hs.515550	0.26	0.69
Hs.128623 || || Transcribed locus, weakly similar to XP_534504.2 PREDICTED: similar to LINE-1 reverse transcriptase homolog [Canis familiaris]	Hs.128623	0.26	1.51
Hs.104829 || || Transcribed locus	Hs.104829	0.27	0.21
Hs.633058 || || Transcribed locus	Hs.633058	0.28	0.09

Genes lower expressed in arthralgia patients (79)

Gene name	Gene ID	Fold Change	q-value (%)
|| || || AA555001::AI791850::AI821608 || || || IMAGE:1029604 || 81385	IMAGE:1029604	0.30	1.51
Hs.417004 || S100A11 || S100 calcium binding protein A11 (calgizzarin)	Hs.417004	0.30	0.00
Hs.471014 || TLN1 || Talin 1	Hs.471014	0.36	1.51
Hs.471955 || RP5-860F19.3 || KIAA1442 protein	Hs.471955	0.36	0.79
Hs.104825 || LOC143244 || **Similar to Eukaryotic translation initiation factor 5A (eIF-5A) (eIF-4D) (Rev-binding factor)	Hs.104825	0.38	0.00
|| || || AI493835 || || || IMAGE:2030430 || 126265	IMAGE:2030430	0.38	0.00
Hs.445233 || || Transcribed locus	Hs.445233	0.39	0.79
|| || || AI203272 || || || IMAGE:1941787 || 111416	IMAGE:1941787	0.41	0.00
|| || || AA535991::AI732345 || || || IMAGE:927639 || 94973	IMAGE:927639	0.41	0.00
Hs.374378 || CKS1B || CDC28 protein kinase regulatory subunit 1B	Hs.374378	0.41	0.00
Hs.469173 || USP39 || Ubiquitin specific peptidase 39	Hs.469173	0.41	3.54
|| || || AI821809::AA662117::AI821062 || || || IMAGE:1187808 || 95378	IMAGE:1187808	0.43	0.21
|| || || AA687438::AA687494::AI821109 || || || IMAGE:1187853 || 116519	IMAGE:1187853	0.43	0.14
Hs.559736 || OLFM4 || Olfactomedin 4	Hs.559736	0.43	3.54
|| || || R78565 || || || IMAGE:144856 || 127778	IMAGE:144856	0.44	0.00
Hs.79625 || C20orf149 || Chromosome 20 open reading frame 149	Hs.79625	0.44	0.00
Hs.500761 || SLC16A3 || Solute carrier family 16 (monocarboxylic acid transporters), member 3	Hs.500761	0.44	0.00
Hs.83347 || AAMP || Angio-associated, migratory cell protein	Hs.83347	0.44	0.00
|| || || AA524273::AI732354 || || || IMAGE:936683 || 28718	IMAGE:936683	0.44	0.00
Hs.517549 || PIB5PA || Phosphatidylinositol (4,5) bisphosphate 5-phosphatase, A	Hs.517549	0.45	0.00
Hs.602915 || || Transcribed locus	Hs.602915	0.45	0.00
Hs.523097 || LOC642592 || Similar to eukaryotic translation initiation factor 5A	Hs.523097	0.46	0.06
Hs.512304 || CCL3L1 || **Chemokine (C-C motif) ligand 3-like 1	Hs.512304	0.46	0.79
Hs.513285 || LOC146439 || Hypothetical LOC146439	Hs.513285	0.46	0.09
|| || || AA235476::AA292540 || || || IMAGE:723866 || 128936	IMAGE:723866	0.46	0.00
|| || || AI265991 || || || IMAGE:1891788 || 73626	IMAGE:1891788	0.46	0.00
Hs.235782 || SLCO4A1 || Solute carrier organic anion transporter family, member 4A1	Hs.235782	0.46	0.00
Hs.632182 || TBC1D10B || TBC1 domain family, member 10B	Hs.632182	0.46	0.40
|| || || AA536013::AI732346 || || || IMAGE:927667 || 116112	IMAGE:927667	0.47	0.06
Hs.502872 || MAP3K11 || Mitogen-activated protein kinase kinase kinase 11	Hs.502872	0.47	0.00
Hs.343380 || CDC42EP2 || CDC42 effector protein (Rho GTPase binding) 2	Hs.343380	0.47	0.12
Hs.122199 || || CDNA clone IMAGE:4829423	Hs.122199	0.47	4.21
Hs.575475 || LOC644150 || Similar to SH3 domain binding protein CR16	Hs.575475	0.47	0.21
Hs.32804 || || Transcribed locus, strongly similar to XP_374188.1 PREDICTED: hypothetical protein XP_374188 [Homo sapiens]	Hs.32804	0.47	0.00

Genes lower expressed in arthralgia patients (79)

Gene name	Gene ID	Fold Change	q-value (%)
Hs.24553 || STRA6 || Stimulated by retinoic acid gene 6 homolog (mouse)	Hs.24553	0.47	0.00
Hs.604701 || || Transcribed locus	Hs.604701	0.47	0.00
|| || || || R27446::R27244 || || || IMAGE:133423 || 28985	IMAGE:133423	0.47	0.00
Hs.107196 || ATPBD4 || ATP binding domain 4	Hs.107196	0.47	0.00
|| || || || AA484302 || || || IMAGE:911483 || 100484	IMAGE:911483	0.47	0.00
Hs.509637 || PLEKHG3 || Pleckstrin homology domain containing, family G (with RhoGef domain) member 3	Hs.509637	0.48	0.28
Hs.1584 || COMP || Cartilage oligomeric matrix protein	Hs.1584	0.48	0.00
Hs.568913 || || Transcribed locus, weakly similar to NP_689672.2 hypothetical protein MGC45438 [Homo sapiens]	Hs.568913	0.48	0.00
|| || || || AI262370 || || || IMAGE:1871101 || 19601	IMAGE:1871101	0.48	0.12
|| || || || AA688217::AI821259 || || || IMAGE:1220441 || 30520	IMAGE:1220441	0.48	0.18
Hs.591835 || LOC389641 || Hypothetical gene supported by AK124295	Hs.591835	0.48	0.00
Hs.548787 || D21S2090E || D21S2090E	Hs.548787	0.48	0.07
Hs.7835 || MRC2 || Mannose receptor, C type 2	Hs.7835	0.48	0.21
|| || || || AI095492 || || || IMAGE:1697156 || 16175	IMAGE:1697156	0.48	0.00
Hs.592115 || TOP3A || Topoisomerase (DNA) III alpha	Hs.592115	0.48	0.00
Hs.633175 || || Transcribed locus	Hs.633175	0.48	0.00
Hs.522639 || SUV39H1 || Suppressor of variegation 3-9 homolog 1 (Drosophila)	Hs.522639	0.48	0.00
Hs.513977 || NT5M || 5',3'-nucleotidase, mitochondrial	Hs.513977	0.49	0.00
|| || || || AA568399 || || || IMAGE:913911 || 6430	IMAGE:913911	0.49	0.12
Hs.239891 || GPR35 || G protein-coupled receptor 35	Hs.239891	0.49	0.06
Hs.37062 || INSL3 || Insulin-like 3 (Leydig cell)	Hs.37062	0.49	0.06
Hs.369252 || SLC22A6 || Solute carrier family 22 (organic anion transporter), member 6	Hs.369252	0.49	0.00
Hs.570637 || || Transcribed locus, strongly similar to XP_520681.1 PREDICTED: similar to prickle-like 2 [Pan troglodytes]	Hs.570637	0.49	0.00
Hs.600218 || || Transcribed locus	Hs.600218	0.49	0.00
Hs.574822 || C20orf55 || Chromosome 20 open reading frame 55	Hs.574822	0.49	0.00
Hs.195633 || LOC644046 || Hypothetical protein LOC644046	Hs.195633	0.49	0.40
Hs.307518 || SLC7A14 || Solute carrier family 7 (cationic amino acid transporter, y+ system), member 14	Hs.307518	0.49	0.00
Hs.457403 || || Transcribed locus	Hs.457403	0.49	3.02
|| || || || AA420876 || || || IMAGE:745690 || 114061	IMAGE:745690	0.49	0.12
Hs.634875 || || CDNA FLJ14459 fis, clone HEMBB1002409	Hs.634875	0.49	0.00
Hs.37860 || KLF1 || Kruppel-like factor 1 (erythroid)	Hs.37860	0.49	1.74
Hs.590921 || CXCL2 || Chemokine (C-X-C motif) ligand 2	Hs.590921	0.50	0.07

Genes lower expressed in arthralgia patients (79)

Gene name	Gene ID	Fold Change	q-value (%)
Hs.380240 || || Transcribed locus, weakly similar to XP_518535.1 PREDICTED: similar to dJ108C2.1.4 (MCM3 minichromosome maintenance deficient 3 (S. cerevisiae), variant 4) [Pan troglodytes]	Hs.380240	0.50	0.00
Hs.1437 || GAA || Glucosidase, alpha; acid (Pompe disease, glycogen storage disease type II)	Hs.1437	0.50	0.07
Hs.523836 || GSTP1 || Glutathione S-transferase pi	Hs.523836	0.50	0.12
Hs.491734 || C1orf92 || Chromosome 1 open reading frame 92	Hs.491734	0.50	4.21

FDR<5% and at least a two-fold change in expression

Selecting genes for validation of microarray data using TLDA

SAM revealed 255 significantly differential expressed genes whose transcript levels differed at least twofold between autoantibody negative healthy individuals and arthralgia patients positive for ACPA and/or RF (supplementary Table S2). In order to narrow down the list of genes for validation, class prediction analysis of microarrays (PAM) was applied using the gene expression data as a training set. Applying a ten-fold cross validation this analysis identified a set of only 17 genes that could correctly classify our controls from autoantibody positive arthralgia patients while only two of the 19 arthralgia patients were classified as controls. Thus, the expression of these 17 genes could predict with a class error rate of only 10.5% if a sample was derived from an autoantibody negative control or an autoantibody positive arthralgia patient at risk of developing RA. An additional PAM analysis was performed in which only data was used of the 6 autoantibody negative controls and the 6 ACPA positive but RF negative persons at risk. Strikingly, the expression of 14 genes could correctly classify these two different groups with a class error rate of 0%. Interestingly, the majority of these 14 genes are interferon-induced and only one gene overlapped with the first PAM analysis. When we performed a PAM analysis between the 6 controls and the 9 ACPA negative but RF positive patients, most of the genes overlapped with the first PAM analysis (6 controls versus 19 autoantibody positive arthralgia patients). Available pre-designed Taqman primers and probes were used to validate gene expression levels of the above PAM analyses. For 10 selected genes no pre-designed assays were available and therefore these genes were excluded from analysis.

We included 20 IFN related genes derived from comparisons between active RA patients and healthy controls performed previously (1). In addition, IFN specific genes IFNα2, IFNβ and IFNγ were included and housekeeping genes GAPDH and 18SRNA were added for normalization. Detailed information for selected target genes based on differential expression between controls and at risk individuals is listed in Table S3.

Genes representative for the observed heterogeneity within the at risk group

were selected from the different clusters displayed in Figure 3. Therefore, per gene cluster patients were re-clustered after which genes were selected that showed the highest variance in expression levels using PAM and SAM analysis. The gene set was supplemented with genes representative for the different biological processes. Detailed information for selected target genes representative for the heterogeneity within the at risk group is listed in Table S4.

In total 87 genes were selected for expression analysis in the total cohort using the Taqman Low Density Array (TLDA) technology which is based on quantitative real-time PCR.

Table S3. Selected target genes based on differential expression between controls and at risk individuals

Symbol	Full name	Reason†
18SRNA	ribosomal 18S-specific	Control
C12orf35	chromosome 12 open reading frame 35	1
C18orf17	chromosome 18 open reading frame 17	3
CCT4	chaperonin containing TCP1, subunit 4 (delta)	1
CD274	CD274 molecule	3
CKS1B	CDC28 protein kinase regulatory subunit 1B	1+2
DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	1
DTX3L	deltex 3-like (Drosophila)	2
EEF1G	eukaryotic translation elongation factor 1 gamma	3
EIF2AK2 =PKR	eukaryotic translation initiation factor 2-alpha kinase 2	3
EPSTI1	epithelial stromal interaction 1 (breast)	2+3
FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64)	3
FLJ31033	hypothetical protein FLJ31033	3
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	Control
GBP1	guanylate binding protein 1, interferon-inducible, 67kDa	3
HERC3	hect domain and RLD 3	1
HLA-G*	histocompatibility antigen, class I, G	4
ID1	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein	1
IFI44L	interferon-induced protein 44-like	3
IFI6	interferon, alpha-inducible protein 6	3
IFIH1 =MDA5	interferon induced with helicase C domain 1	2
IFIT1	interferon-induced protein with tetratricopeptide repeats 1	2
IFIT2	interferon-induced protein with tetratricopeptide repeats 2	3
IFITM1	interferon induced transmembrane protein 1 (9-27)	3
IFNA2	interferon, alpha 2	IFN
IFNB1	interferon, beta 1, fibroblast	IFN
IFNG	interferon, gamma	IFN
IPO7	importin 7	1
IRF2	interferon regulatory factor 2	3
ISG15	ISG15 ubiquitin-like modifier	3
KDELR3*	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3	4
MAT2B	methionine adenosyltransferase II, beta	1
MX1	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	3

OAS1	oligoadenylate synthetase 1, 40/46 kDa	3
OAS2	oligoadenylate synthetase 2, 69/71 kDa	3
OAS3	oligoadenylate synthetase 3, 100 kDa	3
PARP14	poly (ADP-ribose) polymerase family, member 14	2+3
PLSCR1	phospholipid scramblase 1	3
REN*	renin	1
RSAD2	radical S-adenosyl methionine domain containing 2	2
RUFY1	RUN and FYVE domain containing 1	1
SAMD9L	sterile alpha motif domain containing 9-like	2+3
SELL	selectin L (lymphocyte adhesion molecule 1)	1
SERPING1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	3
STAT1	signal transducer and activator of transcription 1, 91kDa	2+3
TNFSF10 =TRAIL	tumor necrosis factor (ligand) superfamily, member 10	4
TPM3	tropomyosin 3	1
TRIM22	tripartite motif-containing 22	3

* Gene expression below detection limit of TLDA analysis. † Reason for selection (See M&M for details): 1. PAM analysis: 6 controls vs. 19 persons at risk; 2. PAM analysis: 6 controls vs. 6 ACPA+/RF- at risk; 3. Active RA vs. healthy controls: IFN-induced genes (1); 4. Explorative SAM analyses between subgroups of patients; IFN: type I/II IFN specific gene; control: gene used for normalization. PCR amplification failed in five cDNA samples (2 controls, 2 autoantibody positive arthralgia patients and 1 RA patient).

Table S4. Selected genes representative for observed heterogeneity within at risk group

Gene Symbol	Full Name	Cluster (Fig. 2)	Used analysis
BAG1	BCL2-associated athanogene	D	SAM
BCL2L1	BCL2-like 1	D	PAM
CCL5	chemokine (C-C motif) ligand 5	B	PAM
CD19	CD19 molecule	A	PAM
CD79A	CD79a molecule	A	PAM
CD79B	CD79b molecule	A	PAM
DARC	Duffy blood group	B	PAM
DEFA3	defensin alpha 3	D	SAM
ERAF	erythroid associated factor	D	PAM
FCGR1A†	Fc fragment of IgG, high affinity Ia, receptor (CD64)	B	PAM
FCRL5	Fc receptor-like 5	A	PAM
FLT3LG	fms-related tyrosine kinase 3 ligand	A	Function
GADD45B	growth arrest and DNA-damage-inducible beta	A	Function
GZMH	granzyme H	B	PAM
HDAC10	histone deacetylase 10	Group III	SAM
IFI27	interferon, alpha-inducible protein 27	B	PAM
IFNAR2	interferon (alpha beta and omega) receptor 2	C	Function
IGKC	immunoglobulin kappa constant, immunoglobulin kappa variable 1-5	A	PAM
IGLL1*	immunoglobulin lambda-like polypeptide 1	A	PAM
IL1A*	interleukin 1 alpha	D	Function
IL32	interleukin 32	A	Function
IL7R	interleukin 7 receptor	A	Function
KLF1	Kruppel-like factor 1 (erythroid)	D	PAM

LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein	A	Function
LTB	lymphotoxin beta (TNF superfamily, member 3)	D	Function
LTF	lactotransferrin	D	PAM
MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa)	B	PAM
MRPL38	mitochondrial ribosomal protein L38	C	PAM
MS4A1	membrane-spanning 4-domains, subfamily A, member 1	A	PAM
MYOM2*	myomesin (M-protein) 2, 165kDa	A	SAM
NKG7	natural killer cell group 7 sequence	B	PAM
PADI2	peptidyl arginine deiminase, type II	D	Function
PBEF1	pre-B-cell colony enhancing factor 1	B	Function
PDK3	pyruvate dehydrogenase kinase isozyme 3	D	PAM
RBM38	RNA binding motif protein 38	D	PAM
RGS18	regulator of G-protein signaling 18	C	SAM
RPL23	ribosomal protein L23	B	PAM
S100A12	S100 calcium binding protein A12	B	Function
S100A8	S100 calcium binding protein A8	B	Function
SELENBP1	selenium binding protein 1	D	PAM
SERPING1†	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	B	PAM
SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	C	PAM
SPP1	secreted phosphoprotein 1 (osteopontin)	D	Function
STAT4	signal transducer and activator of transcription 4	B	Function
TESC	tescalcin	D	SAM
TNFSF7	CD70 molecule	D	SAM
TRGV9	TCR gamma alternate reading frame protein, T cell receptor gamma variable 9	B	PAM
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen	C	SAM

* gene expression below detection limit of TLDA analysis and therefore excluded from analyses. † also present in gene selection based on comparison of autoantibody positive arthralgia patients and healthy controls (Table S3).

Table S5. 52 biomarker gene set

Gene Symbol	Full Name
BAG1	BCL2-associated athanogene
BCL2L1	BCL2-like 1
C18orf17	chromosome 18 open reading frame 17
CCL5	chemokine (C-C motif) ligand 5
CD19	CD19 molecule
CD274	CD274 molecule
CD79A	CD79a molecule
CD79B	CD79b molecule
CKS1B	CDC28 protein kinase regulatory subunit 1B
DARC	Duffy blood group
DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
DEFA3	defensin alpha 3
EPSTI1	epithelial stromal interaction 1 (breast)

ERAF	erythroid associated factor
FCRL5	Fc receptor-like 5
GADD45B	growth arrest and DNA-damage-inducible beta
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GZMH	granzyme H
HDAC10	histone deacetylase 10
HERC3	hect domain and RLD 3
IFI27	interferon, alpha-inducible protein 27
IFI44L	interferon-induced protein 44-like
IFI6	interferon, alpha-inducible protein 6
IFIT1	interferon-induced protein with tetratricopeptide repeats 1
IFNG	interferon, gamma
IL7R	interleukin 7 receptor
IPO7	importin 7
IRF2	interferon regulatory factor 2
ISG15	ISG15 ubiquitin-like modifier
KLF1	Kruppel-like factor 1 (erythroid)
LTB	lymphotoxin beta (TNF superfamily, member 3)
LTF	lactotransferrin
MAT2B	methionine adenosyltransferase II, beta
MRPL38	mitochondrial ribosomal protein L38
MS4A1	membrane-spanning 4-domains, subfamily A, member 1
MX1	myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)
NKG7	natural killer cell group 7 sequence
OAS3	oligoadenylate synthetase 3, 100 kDa
RBM38	RNA binding motif protein 38
RPL23	ribosomal protein L23
RSAD2	radical S-adenosyl methionine domain containing 2
S100A12	S100 calcium binding protein A12
S100A8	S100 calcium binding protein A8
SELENBP1	selenium binding protein 1
SELL	selectin L (lymphocyte adhesion molecule 1)
SERPING1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)
SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1
STAT4	signal transducer and activator of transcription 4
TESC	tescalcin
TNFSF10 =TRAIL	tumor necrosis factor (ligand) superfamily, member 10
TPM3	tropomyosin 3
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen

Supplementary Reference List

(1) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. **Rheumatoid arthritis** subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. **Ann Rheum Dis 2007; 66(8):1008-1014.**

Chapter 3
Genomics for patient subclassification

Chapter 3.1
Rheumatoid Arthritis subtypes identified by genomic profiling of peripheral blood cells: Assignment of a type I interferon signature in a subpopulation of patients

T.C.T.M. van der Pouw Kraan1*, C.A. Wijbrandts2*, L.G.M. van Baarsen1, A.E. Voskuyl3, F. Rustenburg1, J.M. Baggen1, S.M. Ibrahim4, M.Fero5, B.A.C. Dijkmans3, P.P. Tak2, and C.L. Verweij1

1 VU medical center, Dept. of Molecular Cell Biology & Immunology, Amsterdam, the Netherlands. 2 Academic Medical Center/University of Amsterdam, Div. of Clinical Immunology and Rheumatology, Amsterdam, the Netherlands. 3 VU medical center, Dept. of Rheumatology, The Netherlands. 4 University of Rostock, Section of Immunogenetics, Rostock, Germany. 5 Stanford University, Stanford Functional Genomics Facility, Stanford, CA.

** Both authors contributed equally to this study*

Annals of the Rheumatic Diseases 2007 Aug;66(8):1008-14.

Abstract

Objective: Rheumatoid arthritis (RA) is a heterogeneous disease with unknown etiology. Here we aimed to identify peripheral blood gene expression profiles that may distinguish RA subtypes.

Methods: Large-scale expression profiling by cDNA microarrays was performed on peripheral blood from 35 patients and 15 healthy individuals. Differential gene expression was analyzed by Significance Analysis of Microarrays (SAM), followed by Gene Ontology analysis of the significant genes. Gene Set Enrichment Analysis (GSEA) was applied to identify pathways relevant to disease.

Results: We found a remarkably elevated expression of a spectrum of genes involved in immune defense in the peripheral blood of RA patients compared to healthy individuals. SAM analysis revealed a highly significant elevated expression of interferon (IFN) type I regulated genes in RA compared to healthy individuals, which was confirmed by Gene Ontology and Pathway analysis, suggesting that this pathway was activated systemically in RA. A quantitative analysis revealed that increased expression of IFN-response genes was characteristic of approximately half of the patients (IFN^{high} patients). Application of pathway analysis revealed that the IFN^{high} group was largely different from the controls, with evidence for upregulated pathways involved in coagulation and complement cascades, and fatty acid metabolism, while the IFN^{low} group was similar to the controls.

Conclusion: The IFN type I signature defines a subgroup of RA patients, with a distinct biomolecular phenotype, characterized by increased activity of the innate defense system, coagulation and complement cascades, and fatty acid metabolism.

3

Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune disease characterized by chronic inflammation of the joints. There is growing evidence that patients with RA, as defined by the American College of Rheumatology Classification Criteria.(1) represent a highly heterogeneous group. However, the clinical approach to disease classification could erroneously suggest that criteria are applied to classify one disease entity. The heterogeneity of RA is reflected by marked variability in clinical presentation, and the presence of distinct autoantibody specificities, like rheumatoid factor and anti-cyclic citrullinated peptide antibodies (ACPA) in the serum.(2;3)

Disease heterogeneity is also apparent in histological features of the synovium, displaying different complexity levels of lymphoid organization in subsets of patients.(4;5) Moreover, gene expression profiling of synovial tissue from RA and osteoarthritis (OA) patients revealed marked variation in gene expression profiles that allowed us to identify molecularly distinct forms of RA synovium.(6;7)

The wide variation in responsiveness to virtually any treatment in RA is also consistent with the heterogeneous nature of the disease.(8-10) Together these

findings suggest that distinct disease mechanisms are at play in RA pathology. The relative contribution of the different mechanisms may vary among patients and, perhaps, in different stages of disease.

The heterogeneity most likely has its origin in the multifactorial nature of the disease, whereby specific combinations of environmental factors and a varying polygenic background are likely to influence not only susceptibility but also the severity and disease outcome. Findings from genetically identical twins, where the concordance rate is far less than complete, are indicative for a major role of an environmental factor in the risk of developing RA.(11;12)
Given the heterogeneous nature of RA, and its systemic features, we investigated whether this heterogeneity is reflected in peripheral blood cells, because it can be anticipated that (etio)pathogenic events in the host are reflected as phenotypic changes in the host cells. Large-scale gene expression profiling of peripheral blood cells from RA patients could thus provide a molecular portrait that reflects the contributions of diverse cellular responses that are associated with RA in general and with disease subtypes, and thus defines the samples' unique biology.

▪ Material and methods

Patients and controls
Peripheral blood (PB) was obtained in PAXgene RNA isolation tubes (PreAnalytix, GmbH, Germany) from 35 patients. From all 35 patients, 25 used methotrexate (MTX), and 10 patients were MTX and other DMARD naïve. All patients fulfilled the revised American College of Rheumatology 1987 criteria for RA,(1) except for 3 patients in the MTX naive group, which were diagnosed with probable RA with a disease duration of 3-12 months, a mono-arthritis and positive for ACPA. Two patients have been diagnosed with RA after 6 and 12 months. Table 1 summarizes the characteristics of these patients.

Table 1. Patient characteristics.
* available for 23 of 25 patients on MTX treatment. ND indicates not determined

	RA patients on MTX Treatment N=25	RA patients without MTX Treatment N=10
Age (range)	Mean 49 (23-63)	49 (25-67)
Sex: Female, Male	17 F, 8 M	8 F, 2 M
Disease characteristics		
CRP (range)	Mean 19 (3-76)	ND
ESR (range)	Mean 27 (2-82)	41 (13-70)
Rheumatoid factor titer	Mean 114 (1-516)*	123 (12-832)
ACPA titer	Mean 935 (1-6264)*	1210 (1-4904)
Disease duration in months	Mean 113 (8-417)	7 (1-12)
Erosions	N=24	N=1
DAS28	5.5 (3.4-7.2)	5.1 (3.1-7.4)
Medication		
MTX dose (range)	Mean 21 (7.5-30)	0
Prednisone ≤10 mg./day	N=6	N=0

The control group consisted of 15 healthy individuals (9 females, 6 males, mean age: 43 years, ranging from 27-63). In all comparisons mentioned, the groups were age- and sex-matched. All patients and controls gave their informed consent, and the study protocol was approved by the Medical Ethics Committees from the Academic Medical Center and VU medical center.

Sample preparation, labeling and hybridization

This procedure was performed as previously described.(13) In short, total RNA was isolated from PB using the PAXgene RNA isolation kit. Amplified RNA was labeled with aminoallyl-dUTP during cDNA synthesis, followed by chemical coupling of the aminoallyl group to Cy3 or Cy5 for the experimental and reference samples, respectively. The labeled cDNA transcripts were hybridized together on human cDNA microarrays with 42,000 elements, representing ~24,000 genes, generated at Stanford University, as described.(14)

Data filtering and analysis

Data were stored and pre-analyzed in the Stanford Microarray Database (SMD) (15) at http://genome-www.stanford.edu/microarray. Data are expressed as $\log_2$ ratios of fluorescence intensities of the experimental and the common reference sample. Intensity-dependent normalization using local estimation ("Loess") was performed separately on each sector of the array. Spots were included in the analysis when in at least 80% of the microarrays a reliable data point was obtained for that element (defined by a regression correlation coefficient $R > 0.6$ for Cy3 and Cy5 pixel intensities, and a signal intensity of 2.5 times the local background for both Cy3 and Cy5). The use of a common reference allows comparison of the expression levels across all samples.(13) Therefore, the expression levels (as log2 ratio's) were median centered, i.e. each spotted element was expressed relative to the median expression level of that element across all samples. We corrected for array batch differences by applying Single Value Decomposition.(16) Genes represented more than once on the microarrays were averaged in SMD from sequences with the same Unigene identifier.

Statistical analysis

Statistical analysis on microarray data was performed using Significance Analysis of Microarray data (SAM).(17) Genes that were expressed at significantly different levels between patients and controls, defined by a q-value of less than 5%, were analyzed by supervised hierarchical clustering (18) to visualize the correlation of co-expressed genes in Treeview (available at http://rana.lbl.gov/EisenSoftware.htm).

For an interpretation of the biological processes that are represented by the genes that show a significantly different level of expression in RA patients compared to the controls, we applied Gene Ontology analysis in the PANTHER database at http://PANTHER.appliedbiosystem.com.(19) PANTHER uses the binomial statistics

tool to compare our gene list to a reference list (NCBI: Homo sapiens genes) to determine the statistically significant over-representation of functional groups of genes. A Bonferroni correction was applied to adjust for multiple comparisons. P values < 0.05 were considered significant.
For pathway analysis, we used Gene Set Enrichment Analysis (GSEA) (20) at http://www.broad.mit.edu/gsea/. Like SAM, it utilizes data permutation to adjust for multiple testing, indicated by a false discovery rate. A Total of 408 pathways from the Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg) and Biocarta (http://www.biocarta.com) are applied in this analysis. The same gene may be present in more than one pathway or biological process. In addition, we incorporated several IFN-response gene sets from published data.(21;22) A minimal gene set size of 20 genes per pathway was applied, and pathways with a P-value < 0.05 and a false discovery rate (FDR) of < 0.25 were considered significant, according the authors' suggestions.(20) For the comparison of mean gene expression levels in different gene sets, a Student's T test was used.

▪ Results

Gene expression profiling in peripheral blood cells of RA patients
Gene expression profiling of PB cells from 32 RA, 3 probable RA patients, and 15 age and sex-matched healthy controls was performed on microarrays with a complexity of ~20K unique genes (43K elements). Data were analyzed as two class, unpaired data using SAM.(17)] A total of 577 genes, of which 259 were upregulated, and 318 genes downregulated, were selected whose transcript levels were expressed at significantly different levels between the two groups. The siginifcant gene expression differences between RA patients and healthy controls were visualized in a heatmap (Figure 1A).(18)

Genes upregulated in RA.
A global view of the significantly differential expressed genes revealed a prominent cluster of IFN-inducible genes that was upregulated in RA patients. This cluster, highlighted in Figure 1B, contains highly correlated genes such as IFRG28 (28kD interferon responsive protein), IFI35 (interferon-induced protein 35), IFI44L (interferon-induced protein 44-like), IFIT1 (interferon-induced protein with tetratricopeptide repeats 1), IFIT2, IRF2 (interferon regulatory factor 2), IRF7, GIP2 (interferon alpha-inducible protein 2), GIP3, SERPING1 (serine proteinase inhibitor clade G member 1, C1 inhibitor), OAS1 (2'-5'-oligoadenylate synthetase 1), OAS2, MX1 (Myxovirus resistance 1), G1P2/ISG15 (Interferon-induced protein 15), and RSAD2 (radical S-adenosyl methionine domain containing 2).
In addition, all patients showed increased expression of several inflammatory mediators including the chemokines CXCL12, CXCL9, CCL15, CCL19, CCL7, CXCL12, CCL19, CCL7, CXCL3, and CCL8 as well as interleukin (IL)-19 and the S100 family

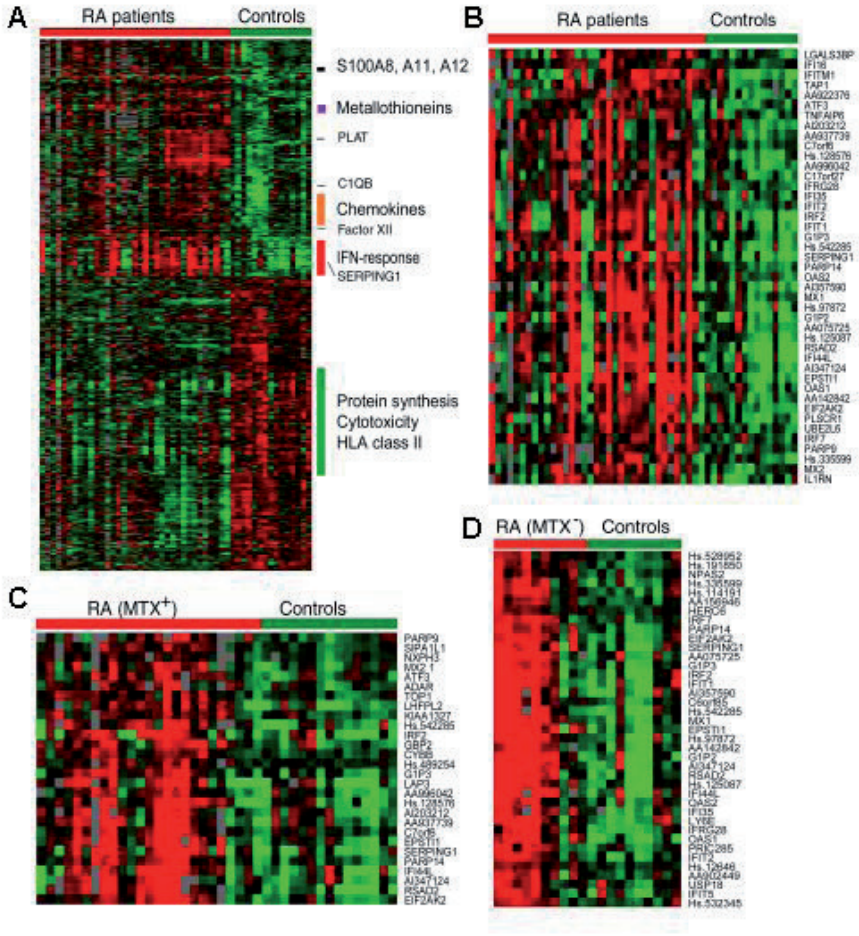


Figure 1

A. Cluster diagram of the expression of 577 significantly different expressed genes in 35 patients and 15 healthy individuals. Genes are organized by hierarchical clustering based on overall similarity in expression patterns.(18) Red represents relative expression greater than the median expression level across all samples, and green represents an expression level lower than the median. Black indicates intermediate expression. Grey indicates missing data. Colored bars to the right identify the locations of a category of clustered genes with a correlated expression profile and related function.

B. Representation of the IFN-response gene cluster with an enhanced expression in the RA patient group.
An expanded view is shown of the genes in the IFN-response cluster of Figure 1A. Genes are either known genes with a unigene symbol that is characteristic for the defined gene cluster, or genes are unknown, indicated by an accession number or unigene cluster ID.

C and D. The IFN-response program is present in RA patients irrespective of treatment.
Representation of genes that are expressed at significantly different levels between RA patients on MTX treatment (C) or patients without MTX treatment (D) and age-and sex matched healthy controls. A selection of genes with a correlated expression profile is shown that are indicative for an IFN-response program.

proteins S100 calcium-binding protein A8 (S100A8), S100A11, and S100A12. Other genes that were upregulated in RA patients were members of the antioxidant metallothionein family and the anti-inflammatory IL-1 receptor antagonist.

Genes downregulated in RA.
Genes that showed a lower expression in RA included: CD3 zeta, TCR beta chain, TARP/TCRgammaV9, granzyme M, runx3, and KLRB1, which are involved in cytotoxic functions, and many other genes with unknown function.

Gene Ontology analysis of genes with significant differential expression in RA
To systematically categorize the 577 genes with significant differential expression into functional groups we used the PANTHER database consisting of a large collection of protein families that have been subdivided into functionally related subfamilies.(19) The differentially upregulated genes represented 7 significant functional biological processes (Table 2). There were no significant downregulated ontology groups. The "Immunity and Defense" ontology group represents a broad composite family that consists of more specified ontology subgroups. Within these subgroups the most significant upregulated process that distinguished RA patients from controls was "Interferon-Mediated Immunity"

Table 2.
Ontology analysis of the genes that were expressed at significantly different levels between all 35 patients, or the subgroups consisting of 20 IFN^{high} and 15 IFN^{low} patients, and healthy controls. ns indicates non-significant. Processes with a P-value < 0.05 were considered significant (with Bonferroni correction).

	All patients	**IFN^{high}**	**IFN^{low}**
Biological Process	p-value	p-value	p-value
Immunity and defense	1.61E-11	3.41E-15	ns
Interferon-mediated immunity	1.09E-06	3.38E-08	ns
Macrophage-mediated immunity	6.83E-05	1.27E-04	ns
Lipid and fatty acid transport	1.54E-04	ns	ns
Transport	4.82E-04	3.33E-02	ns
Cytokine/chemokine mediated immunity	1.63E-03	2.61E-05	ns
Ligand-mediated signaling	3.52E-03	1.92E-02	ns
Cell motility	ns	4.24E-03	ns
Blood clotting	ns	5.30E-03	ns
Oncogenesis	ns	1.04E-02	ns
Cell structure and motility	ns	1.55E-02	ns
Inhibition of apoptosis	ns	1.66E-02	ns
Signal transduction	ns	3.21E-02	ns
Proteolysis	ns	3.97E-02	ns

Pathway analysis
In addition, we performed Gene Set Enrichment Analysis (GSEA) (20) to identify

pathways relevant to RA. In contrast to ontology analysis this algorithm is based on the usage of all available gene expression data and derives it power from the analysis of sets of genes that are coordinately regulated in a defined biological process or pathway, while it uses data permutation to adjust for multiple testing. In addition to the intrinsic GSEA pathway gene sets, we included previously reported IFN-response sets in our analysis.(21;22) The results revealed that besides the five GSEA intrinsic gene sets (Table 3), the previously described type I IFN-induced genes by Baechler et al.(22) (in their supplementary data), and the IFNα-induced genes were both significantly increased in RA patients.

Table 3

Pathways which are overexpressed in all patients, and in the subgroups of IFN^{high} and IFN^{low} patients, all compared to healthy controls, analysed by Gene Set Enrichment Analysis. ns indicates nonsignificant. Pathways are shown with a P-value < 0.05, and a FDR value < 0.25.

		All patients	IFN^{high}	IFN^{low}
Pathway	# genes	p-value	p-value	p-value
Type I IFN-induced genes (22)	83	0.004	<0.001	ns
IFNα-induced (21)	36	0.022	<0.001	ns
IFNβ-induced (21)	51	ns	<0.001	ns
Cytokine-cytokine receptor interaction	129	0.006	0.005	ns
Neuroactive ligand-receptor interaction	90	0.009	ns	<0.001
Jak-STAT signaling pathway	86	0.016	0.008	ns
Complement and coagulation cascades	44	0.025	0.009	ns
Fatty acid metabolism	41	0.027	0.015	ns
IFNγ-induced (21)	34	ns	0.019	ns

IFN-induced genes in RA

We confirmed expression of key genes of the IFN pathway, RSAD2 and G1P2, in all samples by real-time PCR, which showed a high correlation with the microarray data ($r = 0.78$ and 0.87 respectively, $P < 0.0001$ in both cases, data not shown). To rule out an effect of MTX treatment on the IFN-induced genes, we made a comparison of patients with (n=25) and without MTX treatment (n=10) with the appropriate age and sex-matched controls. SAM revealed that both groups of patients showed a prominent IFN-induced gene expression signature (Figure 1C and D). Thus, the IFN expression signature was present in RA patients irrespective of MTX treatment.

The expression profiles of the three probable RA patients, within the MTX naive group, did not differ from the other MTX-naive patients. Meanwhile, two of the probable RA patients have been diagnosed with "definite" RA at 6 and 12 months after Paxgene blood sampling, respectively, suggesting that the RA signature is present in the blood prior to diagnosis.

Selective upregulation of type I IFN-response genes in RA

Type I IFNs are mainly produced directly after viral infection whereas type II IFNs are secondary produced by activated T- and natural killer (NK) cells. Type I and type II IFN response programs share many of their genes. To disclose information on the inducing type of IFN, we obtained a specific type I IFN- and type II IFN-response gene set (21) (Supplementary Table 1). The type I IFN-response set consists of 5 genes that respond to both IFNα and IFNβ, but not to IFNγ. The type II IFN-response set consists of 13 genes responding specifically to IFNy.(21) To investigate the relative contribution of either gene set to the RA gene expression profile we calculated for each gene set the mean gene expression level (log2 ratio) per patient and healthy control and compared the two groups with each other (Figure 2). This analysis showed that the mean gene expression level of the type I IFN gene set was significantly higher in the RA patient group (P = 0.0004), whereas the mean gene expression level of the type II IFN genes was similar between patients and controls. Hence, these findings provide evidence that type I IFNs rather than type II IFNs are responsible for the increased expression of IFN-induced genes.

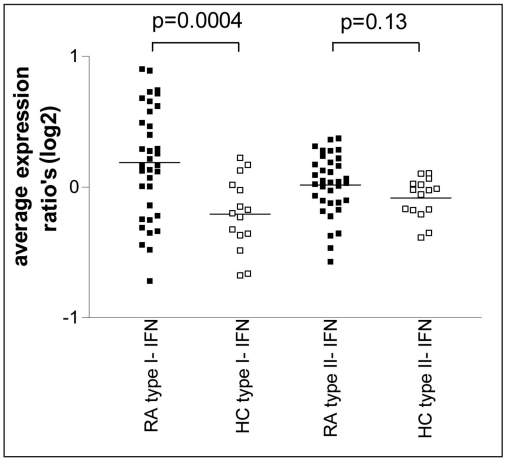


Figure 2
Type I IFN-induced genes are overexpressed in RA.
Each square indicates the mean expression levels of genes known to be specifically induced either by type I (13 genes) or type II IFN (5 genes) per individual patient or healthy control (HC). These genes are extracted from the gene sets used for pathway analysis in Table 3.(21)

The IFN signature defines a subgroup of RA patients.

Consistent with the heterogeneous nature of RA we observed that the IFN-response showed a large variation between RA patients (Figures 2 and 3). To obtain more insight into the differential expression of IFN-induced genes in individual patients, we calculated for each individual, the average expression of the IFN-cluster genes, that were upregulated in RA patients, as described in Figure 1B. Next we defined which patients show an altered IFN-response, by calculating the 95 % limits of the controls (normal values, defined as the mean expression of the 43 IFN genes, plus or minus 1.96 times the standard deviation). We identified 20 patients with an average expression level above normal values, further defined as the IFN^{high} group, while the remainder of the patients, with an expression level equal to controls, were defined as IFN^{low} (Figure 3).

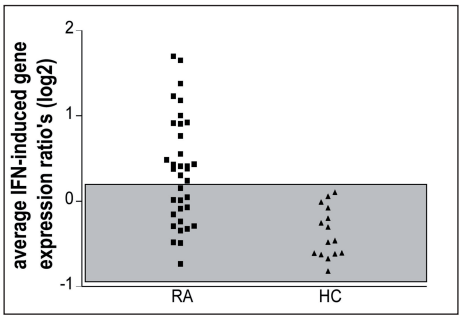


Figure 3
A subgroup of RA patients shows increased expression of IFN-response genes (IFN^{high}).
Each square represents a single individual with the average expression ratio's of all 43 IFN-response genes, that formed a distinct cluster in Figure 1A and B. The shaded box indicates the normal range within the 95% confidence limits. RA patients outside the shaded box are defined as IFN^{high} patients.

Distinct characteristics of the IFN^{high} group

To further characterize the IFN^{high} group, we performed SAM analysis, which revealed that 484 genes were upregulated in IFN^{high} patients, compared to the healthy controls, while 229 genes were downregulated. The same analysis for the IFN^{low} patients revealed only 57 upregulated genes and 93 downregulated genes. These data indicate that within the RA patients, the patients with an IFN signature represent the most distinct group compared to normal controls.
When we applied Gene Ontology analysis, we found that nearly all of the processes that were identified as upregulated in the whole RA group, were also upregulated in the IFN^{high} group. Moreover, an additional group of 10 biological processes were selectively upregulated in the IFN^{high} group. (Table 2). No downregulated processes were identified.
Gene Ontology analysis of the IFN^{low} group revealed no significant down or upregulated processes.

In accordance with Gene Ontology analysis, pathway analysis by GSEA revealed that the IFN^{high} patients were responsible for the upregulated pathways in the overall RA group (Table 3). This was particularly clear for the IFN-type I induced gene sets, complement, and coagulation cascades. On the other hand, the IFN^{low} group was associated with increased expression of the "Neuroactive ligand-receptor interaction" pathway. We did not identify any downregulated pathwways in either group of RA patients. Overall, these analyses indicate that within the whole group of patients, the IFN^{high} group is more distinct from controls than the IFN^{low} group.
The molecular stratification of RA was not associated with clinical parameters that are described in Table 1.

▪ Discussion

Since RA is a systemic disease, several investigators addressed the question whether disease characteristics are reflected by changes in gene expression levels in PB cells. Whereas these studies provided insight into the genes that were differentially expressed between RA patients and healthy controls, the issue of transcript-based

disease heterogeneity has not been addressed so far, except for a comparison between recent onset arthritis and longstanding disease.(23)

Using large-scale gene expression profiling, we identified a large number of genes, including genes involved in the immune/inflammatory response, such as the previously described calcium-binding proteins S100A8, S100A12, and IL1RA.(24;25) Pathway level analysis was used to classify gene expression data in biological processes and pathways. The clear induction of IFN-response genes in RA patients prompted us to incorporate several IFN-response gene sets from published data (21;22) in the analysis. This analysis revealed that the type I IFN-mediated immunity was the most significantly upregulated pathway in RA patients, independent of MTX treatment. Albeit, that inclusion of the type I IFN gene set is a biased decision, this approach provides a method to demonstrate the significance of the type I IFN response program in RA.

Most interestingly, our analysis revealed a striking heterogeneity between RA patients based on the differential expression of genes involved in the innate defense system, in particular the type I IFN system. These findings suggest that different pathogenic mechanisms may contribute to the disease. The IFN^{high} group was further characterized by gene sets reflecting increased activity of complement and coagulation cascades.(26) Next to complement activation, the other pathways associated with the IFN type I signature, such as "Fatty acid metabolism" and "Coagulation" may all contribute to the increased risk for cardiovascular disease in a subgroup of RA patients.(27)
The most significant genes from the complement and coagulation pathway are indicated in Fig 1, including Complement subcomponent C1q chain B (C1QB), Coagulation factor XII (F12), tissue plasminogen activator (PLAT) and Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1 (SERPING1). These genes are involved in activation as well as inhibitory components of the pathways.
Upregulation of IFN-induced genes has also been observed in PB cells of (a subset of) patients with other autoimmune diseases like systemic lupus erythematosus (SLE), (22;28) systemic sclerosis (SSc),(29) multiple sclerosis (MS),(13) and in tissues from patients with Sjögren's syndrome (SS),(30) type I diabetes mellitus,(31) and dermatomyositis.(32) These findings suggest that an activated IFN gene expression program is a common hallmark in chronic autoimmune diseases.

Type I IFNs, which are the early mediators of the innate immune response that influence the adaptive immune response through direct and indirect actions on dendritic cells (DC), T- and B cells, and NK cells, could impact the initiation or amplification of autoimmunity and tissue damage through their diverse and broad actions on almost every cell type and promotion of T helper 1 responses.(33) This appears to be the case for SLE, but for RA both clinical and pathophysiological data

have suggested that tumor necrosis factor alpha (TNFα) rather than type I IFN is essential for persistence of the disease. Hence, it is believed that mutually exclusive cytokine expression patterns are characteristic for distinct autoimmune diseases. However, since we observed an IFN type I signature in the PB of a subgroup of RA patients this could mean that cytokine profiles are a patient-specific rather than a disease-specific phenomenon.

In SLE patients, the IFN signature is related to disease severity.(22) It is at present unclear what the role of type I IFNs in RA pathogenesis could be. In analogy to SSc(29) and MS,(13) no clinical associations were found for RA so far. We have previously suggested that IFN/STAT-1 activation in RA synovium could be a reactive attempt to limit inflammation.(34) This suggestion was recently supported by a study showing that IFNβ deficiency could prolong experimental arthritis and resulted in increased activation of FLS in vitro.(35) In addition, IFNβ-competent fibroblasts were able to ameliorate arthritis in IFNβ-deficient recipients. It should be noted, however, that systemic administration of IFNβ was unsuccessful in the treatment of RA, which may be due to pharmacokinetic issues.(36)

Concerning the origin of type I IFNs, infectious and endogenous agents, such as viruses, bacteria, unmethylated CpG DNA, single- or double stranded RNA, heat shock protein 60, or fibrinogen fragments could all be proximal mediators of type I IFN production and thus lead to the more downstream activation program. In sera from SLE patients, IFNα levels correlate with IFN-response gene expression levels of PB cells.(37) It has been demonstrated for SLE and SS that immune complexes of autoantibodies and DNA- or RNA-containing autoantigens, can induce type I IFN production by PB plasmacytoid DCs (pDCs).(28;30;38;39) This response is dependent on interaction with FcγRIIa and Toll-like receptors.(30;39) Further studies need to determine whether the increased type I IFN-response genes in RA are the result of endogenous or infectious factors.
Besides a role for peripheral blood cells as producers of type I IFNs, cells at the site of inflammation may also be responsible for production. Cells with morphological and phenotypic characteristics of pDC were shown to infiltrate skin lesions in SLE and actively produce type 1 IFN locally.(40) In patients with SS, numerous IFNα -producing cells were detected in the affected salivary gland biopsies.(30) In RA, IFNβ protein has been detected in the synovium.(41) Moreover, fibroblast-like synoviocytes (FLS) were responsible for increased levels of IFNβ in the RA synovium.(42) The endogenous TLR3 ligand, dsRNA derived from necrotic synovial fluid cells, has been shown to stimulate the production of IFNβ in RA FLS.(43)

In conclusion, we demonstrated that genomic profiling powers disease subclassification and has led to the identification of subgroups of patients, based on differential expression of genes involved in non-specific immunity.

Funding:

Supported in part by the EURO-RA Marie Curie Trainings network, the European Community's FP6 integrated program funding (AUTOCURE), the Innovation Oriented research Program (IOP) on Genomics, and the Centre for Medical Systems Biology (a center of excellence approved by the Netherlands Genomics Initiative/ Netherlands Organization for Scientific Research). This publication reflects only the author's views. The European Community is not liable for any use that may be made of the information herein.
These sponsors had no involvement in the study design, analysis or interpretation of the data and publications.

Reference List

(1) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988; 31(3):315-324.

(2) Zendman AJ, van Venrooij WJ, Pruijn GJ. Use and significance of anti-CCP autoantibodies in rheumatoid arthritis. Rheumatology (Oxford) 2006; 45(1):20-25.

(3) Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, Fries JF et al. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. Arthritis Rheum 2005; 52(9):2645-2655.

(4) Tak PP, Smeets TJ, Daha MR, Kluin PM, Meijers KA, Brand R et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. Arthritis Rheum 1997; 40(2):217-225.

(5) Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM et al. Lymphoid neogenesis in rheumatoid synovitis. J Immunol 2001; 167(2):1072-1080.

(6) van der Pouw Kraan TC, van Gaalen FA, Kasperkovitz PV, Verbeet NL, Smeets TJ, Kraan MC et al. Rheumatoid arthritis is a heterogeneous disease: evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. Arthritis Rheum 2003; 48(8):2132-2145.

(7) van der Pouw Kraan TC, van Gaalen FA, Huizinga TW, Pieterman E, Breedveld FC, Verweij CL. Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair. Genes Immun 2003; 4(3):187-196.

(8) Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group 1. N Engl J Med 2000; 343(22):1594-1602.

(9) Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S et al. Treatment of rheumatoid arthritis by selective inhibition of T-cell activation with fusion protein CTLA4Ig. N Engl J Med 2003; 349(20):1907-1915.

(10) Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004; 350(25):2572-2581.

(11) Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 1993; 32(10):903-907.

(12) Deighton CM, Wentzel J, Cavanagh G, Roberts DF, Walker DJ. Contribution of inherited factors to rheumatoid arthritis. Ann Rheum Dis 1992; 51(2):182-185.

(13) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006; 7(6):522-531.

(14) Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA et al. Molecular portraits of human breast tumours. Nature 2000; 406(6797):747-752.

(15) Ball CA, Awad IA, Demeter J, Gollub J, Hebert JM, Hernandez-Boussard T et al. The Stanford Microarray Database accommodates additional microarray platforms and data formats. Nucleic Acids Res 2005; 33(Database issue):D580-D582.

(16) Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, O'Connell JX et al. Molecular characterisation of soft tissue tumours: a gene expression study. Lancet 2002; 359(9314):1301-1307.

(17) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-5121.

(18) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-14868.

(19) Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 2005; 33(Database issue):D284-D288.

(20) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102(43):15545-15550.

(21) Der SD, Zhou A, Williams BR, Silverman RH. Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. Proc Natl Acad Sci U S A 1998; 95(26):15623-15628.

(22) Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2003; 100(5):2610-2615.

(23) Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J et al. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. Ann Rheum Dis 2004; 63(11):1387-1392.

(24) Batliwalla FM, Baechler EC, Xiao X, Li W, Balasubramanian S, Khalili H et al. Peripheral blood gene expression profiling in rheumatoid arthritis. Genes Immun 2005; 6(5):388-397.

(25) Bovin LF, Rieneck K, Workman C, Nielsen H, Sorensen SF, Skjodt H et al. Blood cell gene expression profiling in rheumatoid arthritis. Discriminative genes and effect of rheumatoid factor. Immunol Lett 2004; 93(2-3):217-226.

(26) Wouters D, Wiessenberg HD, Hart M, Bruins P, Voskuyl A, Daha MR et al. Complexes between C1q and C3 or C4: novel and specific markers for classical complement pathway activation. J Immunol Methods 2005; 298(1-2):35-45.

(27) Solomon DH, Goodson NJ, Katz JN, Weinblatt ME, Avorn J, Setoguchi S et al. Patterns of cardiovascular risk in rheumatoid arthritis. Ann Rheum Dis 2006; 65(12):1608-1612.

(28) Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. Arthritis Rheum 2005; 52(5):1491-1503.

(29) Tan FK, Zhou X, Mayes MD, Gourh P, Guo X, Marcum C et al. Signatures of differentially regulated interferon gene expression and vasculotrophism in the peripheral blood cells of systemic sclerosis patients. Rheumatology (Oxford) 2006; 45(6):694-702.

(30) Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML et al. Activation of the type I interferon system in primary Sjogren's syndrome: a possible etiopathogenic mechanism. Arthritis Rheum 2005; 52(4):1185-1195.

(31) Huang X, Yuang J, Goddard A, Foulis A, James RF, Lernmark A et al. Interferon expression in the pancreases of patients with type I diabetes. Diabetes 1995; 44(6):658-664.

(32) Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005; 57(5):664-678.

(33) Biron CA. Interferons alpha and beta as immune regulators--a new look. Immunity 2001; 14(6):661-664.

(34) Tak PP. IFN-beta in rheumatoid arthritis. Front Biosci 2004; 9:3242-3247.

(35) Treschow AP, Teige I, Nandakumar KS, Holmdahl R, Issazadeh-Navikas S. Stromal cells and osteoclasts are responsible for exacerbated collagen-induced arthritis in interferon-beta-deficient mice. Arthritis Rheum 2005; 52(12):3739-3748.

(36) van Holten J, Pavelka K, Vencovsky J, Stahl H, Rozman B, Genovese M et al. A multicentre, randomised, double blind, placebo controlled phase II study of subcutaneous interferon beta-1a in the treatment of patients with active rheumatoid

arthritis. Ann Rheum Dis 2005; 64(1):64-69.

(37) Hua J, Kirou K, Lee C, Crow MK. Functional assay of type I interferon in systemic lupus erythematosus plasma and association with anti-RNA binding protein autoantibodies. Arthritis Rheum 2006; 54(6):1906-1916.

(38) Lovgren T, Eloranta ML, Bave U, Alm GV, Ronnblom L. Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG. Arthritis Rheum 2004; 50(6):1861-1872.

(39) Lovgren T, Eloranta ML, Kastner B, Wahren-Herlenius M, Alm GV, Ronnblom L. Induction of interferon-alpha by immune complexes or liposomes containing systemic lupus erythematosus autoantigen- and Sjogren's syndrome autoantigen-associated RNA. Arthritis Rheum 2006; 54(6):1917-1927.

(40) Farkas L, Beiske K, Lund-Johansen F, Brandtzaeg P, Jahnsen FL. Plasmacytoid dendritic cells (natural interferon- alpha/beta-producing cells) accumulate in cutaneous lupus erythematosus lesions. Am J Pathol 2001; 159(1):237-243.

(41) Holten JV, Smeets TJ, Blankert P, Tak PP. Expression of interferon {beta} (IFN-{beta}) in synovial tissue from rheumatoid arthritis patients compared to osteoarthritis and reactive arthritis patients. Ann Rheum Dis 2005.

(42) Pilling D, Akbar AN, Girdlestone J, Orteu CH, Borthwick NJ, Amft N et al. Interferon-beta mediates stromal cell rescue of T cells from apoptosis. Eur J Immunol 1999; 29(3):1041-1050.

(43) Brentano F, Schorr O, Gay RE, Gay S, Kyburz D. RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via toll-like receptor 3. Arthritis Rheum 2005; 52(9):2656-2665.

▪ Supplementary Table 1

IFN type I and type II –induced Gene sets.
Genes were selected from the previously described IFN-induced genes [20]

Type II IFN-induced genes

IL15RA
BAK1
VAT1
SF3A1
ICAM1

Type I IFN-induced genes

DDX17
OAS1
IFIT1
MX1
TRIM22
PDXK
MX2
G1P3
POLR2B
CD164
SMAD4
OAS2
IRF2

Chapter 3.2

Expression of a pathogen-response program in peripheral blood cells defines a subgroup of Rheumatoid Arthritis patients

3

Lisa G.M. van Baarsen[1,2*], Tineke C.T.M. van der Pouw Kraan[1*], Carla A. Wijbrandts[3], Alexandre E. Voskuyl[4], François Rustenburg[1,2,] Josefien M. Baggen[1], Ben A.C. Dijkmans[4], Paul P. Tak[3], Cornelis L. Verwei[1,2]

[1] Dept. of Molecular Cell Biology & Immunology, [2]Dept. of Pathology, VU University medical center, Amsterdam, the Netherlands; [3]Div. of Clinical Immunology and Rheumatology, Academic Medical Center/University of Amsterdam, the Netherlands, [4]Dept. of Rheumatology, VU University medical center, Amsterdam, The Netherlands

** Both authors contributed equally to this study*

Genes & Immunity 2008 Jan;9(1):16-22

Abstract

Rheumatoid arthritis (RA) is a heterogeneous disease with unknown etiology. Here we aimed to distinguish RA subtypes based on peripheral blood gene expression profiles in comparison with a pathogen-response transcriptional program. Peripheral blood was obtained from 35 RA patients and 15 healthy individuals. For expression profiling we used DNA microarrays. A combined cluster analysis of RA and control samples together with samples from a viral infection model revealed that the gene expression profile of a subgroup of RA patients (RA^A) was reminiscent to that of poxvirus-infected macaques. Statistical analysis, followed by Gene Ontology analysis of the RA^A patients confirmed that these patients form a distinct group, with activation of several host defense mechanisms that resemble a common host-pathogen response. Analysis of the promoter region of genes that were over-expressed in the RA^A patients, revealed an enrichment of transcription factor binding sites for NFκB and interferon-activated transcription factors. Moreover, this subgroup of RA patients expressed significantly increased titers of anti-cyclic citrullinated peptide antibodies.
We conclude that activation of a host-pathogen response defines a subgroup of RA patients characterized by increased autoreactivity against citrullinated proteins.

3

Introduction

RA is a systemic autoimmune disease characterized by chronic inflammation of the joints. The differential expression of the disease is reflected by the marked differences in the abundance and distribution of T and B cells in the rheumatoid synovium, ranging from a scarce and diffuse distribution to the abundant presence of mononuclear cells, which are organized in germinal center-like structures (1;2).The heterogeneous character of RA is further illustrated by the differential responsiveness to treatment (3-5) and the presence of distinct autoantibody specificities, like rheumatoid factor and anti-citrullinated peptide antibodies (ACPA) in the serum (6;7) .
The existence of such molecular heterogeneity in RA synovium fits a model proposed by Firestein and Zvaifler (8), who suggested two independent processes, an immune-mediated and a stromal cell-driven form, which might drive destruction of bone and cartilage.

Development of different types of RA may result from specific combinations of environmental factor(s) and a varying polygenic background. Findings from genetically identical twins, where the concordance rate is far less than complete, are indicative for a major role of an environmental factor in the risk of developing RA (9;10). From animal models it is clear that an exogenous challenge, like infectious agents may induce arthritis. In man, microbial infections are known to contribute to arthritis (11;12). Combined analysis of the EBV-, CMV-, and Parvo virus B19-data

demonstrated that virus infections are far more common in RA or psoriatic arthritis than in reactive arthritis (13). Despite this knowledge the environmental factor(s) that contribute to RA remain to be determined.
Upon infection with a pathogen, cells undergo a marked reprogramming of their transcriptome, providing a direct link between infection and the transcriptional profile of the host. Surprisingly, different pathogens induce a remarkable overlap of transcriptional changes in the host, leading to the concept of a common pathogen response (14). Therefore, analysis of the host transcriptome overcomes the problematic detection of multiple distinct microorganisms, in case no single microorganism but a variety of microorganisms is involved. Large-scale gene expression profiling using DNA microarray technology in blood cells from patients with RA and healthy controls would enable us to detect a molecular portrait representing the contributions and interactions of numerous distinct cells and diverse factors that may be associated with disease. Consequently, if there is a pathogen involved in RA, a systematic analysis of the gene expression profile should reveal similarities with a pathogen response program.
We previously provided evidence for an activated IFN-type I response gene expression program in a subset of patients with RA (15). This could be indicative for a viral contribution. Here we created a framework for evaluating the relationship between a viral response program and RA in more detail. Therefore, we compared the gene expression profiles of RA patients with the profiles of an *in vivo* virus-infection model that was examined on the same microarray platform, allowing a direct comparison.

Material and methods

Patients and controls
Peripheral blood (PB) was obtained in PAXgene RNA isolation tubes (PreAnalytix, GmbH, Germany) from 35 patients, as described previously (15). From all 35 patients, 25 used methotrexate, and 10 patients were DMARD naive. The control group consisted of 15 age- and sex-matched healthy individuals, as described. (15). All patients and controls gave their informed consent and the Medical Ethics Committees from AMC and VUmc approved the study protocol.

ACPA measurements
ACPA titers were measured by ELISA (Euro Diagnostica No. RA-96RT, Arnhem, the Netherlands) with a lower detection limit of <1 kAU/ml.

Sample preparation, labeling and hybridization
Total RNA was isolated from PB using the PAXgene RNA isolation kit according to the manufacturers' instructions including a DNAse (Qiagen, Venlo, The Netherlands) step to remove genomic DNA. 1 μg of total RNA was linear amplified using the

MessageAmp™ kit from Ambion. An amplified common reference sample was used for all array hybridizations, consisting of a mixture of mRNAs isolated from 11 different cell lines supplemented with RNA from synovial tissue, synovial fibroblasts, and activated PB mononuclear cells (PBMCs).
5 µg of amplified RNA was labeled with aminoallyl-dUTP during cDNA synthesis by reverse transcriptase, followed by chemical coupling of the aminoallyl group to Cy3 or Cy5 for the experimental and reference samples, respectively. The Cy3- and Cy5-labelled cDNA transcripts were hybridized together on 43K human cDNA microarrays (Stanford University) at 65°C for 14 to 18 hours. After washing, arrays were scanned using a G2505B microarray scanner (Agilent Technologies, Amstelveen, the Netherlands). Feature extraction analysis was performed using GenepixPro version 3.0.6.90 (Axon instruments, Union City, CA).

Data filtering and analysis
Data were stored and pre-analyzed in the Stanford Microarray Database (SMD) (16) at http://genome-www.stanford.edu/microarray. Data are expressed as $\log_2$ ratios of fluorescence intensities of the experimental and the common reference sample. Intensity-dependent normalization using local estimation ("Loess") was performed separately on each sector of the array. In addition, 'scale' normalization was applied, which makes the data more comparable across different arrays. Spots were included in the analysis when in at least 80% of the microarrays a reliable data point was obtained for that element (defined by a regression correlation coefficient R>0.6 assuring a linear fit between the Cy3 and Cy5 pixel intensities, and a signal intensity of 2.5 times the local background for both Cy3 and Cy5). The use of a common reference sample allows comparison of the relative expression levels across all samples. Therefore, the expression levels (as log2 ratio's) were median centered, i.e. each spotted element was expressed relative to the median expression level of that element across all samples. We corrected for array batch differences by applying Single Value Decomposition (17). Genes represented more than once on the microarrays were averaged in SMD from sequences with the same Unigene identifier.

Application of published data from the nonhuman primate model of Smallpox
A genome-wide investigation of poxvirus infection in macaques was performed by Rubins et al. (18). In this study, PBMCs were obtained from macaques before and after infection with smallpox virus. Gene expression data was obtained using the same microarray platform and data was filtered as described above. The genes in the smallpox dataset were centered separately to remove the bias of a different reference sample. Two-way hierarchical clustering (19) of macaque and human samples was used to subclassify RA patients and visualize the correlation of co-expressed genes in Treeview (available at http://rana.lbl.gov/EisenSoftware.htm).

Statistical analysis
Statistical analysis on microarray data was performed using Significance Analysis of Microarrays (SAM) (20).
For an interpretation of the biological processes that are represented by the genes that show a significantly different level of expression in RA patients compared to the controls, we used Gene Ontology analysis using the PANTHER (**P**rotein **AN**alysis **TH**rough **E**volutionary **R**elationships) Classification System at http://PANTHER.appliedbiosystem.com (21). PANTHER uses the binomial statistics tool to compare our gene list to a reference list (NCBI: Homo sapiens genes) to determine the statistically significant over-representation of functional groups of genes. A Bonferroni correction was applied to adjust for multiple testing.

▪ Results

Identification of a virus response signature in a non-human primate model.
In order to create a framework, which allows a direct comparison of the RA gene expression profile with that of a physiological viral response program we made use of available *in vivo* data on smallpox virus infected macaques (18). This non-human primate model for smallpox infection provides a detailed view of the *in vivo* host transcriptional program in peripheral blood (PB) after infection. Because the microarrays used in this study were of the same platform as the ones used in the present study, we could make a direct comparison. First, we extracted microarray data from 9 macaques before, and 2 or 3 days post-infection from the public access website of the Stanford Microarray Database (SMD) (16). Next, we identified the genes that significantly changed in expression in peripheral blood cells of infected macaques by Significance Analysis of Microarrays (SAM) (20). At a false discovery rate of 5%, 911 genes showed a significant change in expression after smallpox infection.

The PB gene expression profile of a subgroup of patients with RA is reminiscent of a viral response program.
We previously analyzed peripheral blood gene expression profiles from RA patients, and compared their profiles with healthy controls using microarrays with a complexity of ~26K unique genes (total 43K elements) (15).
Expression of the viral response genes was examined in a combined data set including gene expression data from all RA patients and healthy controls, uninfected and infected macaques. Two-way hierarchical clustering clearly separated the profiles of the macaques before and after viral infection (Figure 1), whereas almost all of the healthy controls clustered in the same group as the uninfected animals. Most interestingly, a total of 23 of the 35 RA patients clustered together with the infected macaques.

The subgroup of RA patients that clustered together with the virus-infected primates, was designated the RA^A subgroup. The patients that co-clustered with the non-infected primates and healthy individuals were referred to as the RA^B subgroup (Table 1).

Table 1. Patient characteristics.

	RA^A N=23	RA^B N=12
Age (mean)	49	49
Sex: Female, Male	17 F, 6 M	8 F, 4 M
Disease characteristics:		
ESR (mean)	31	29
Rheumatoid factor titer (mean)	129 *	95
ACPA titer (mean)	1372*	398**
Disease duration in months (mean)	57	132
Erosions	N=14	N=11
DAS28 (mean)	5.4	5.3
Medication:		
MTX dose (mean)	14	18
Prednisone ≤10 mg./day	N=6	N=0

Patients were subdivided in RA^A and RA^B, based on two-way co-clustering with virus-infected macaques, as presented in Figure 1. * available for 21 of 23 patients, ** significantly different between RA^A and RA^B (p=0.03)

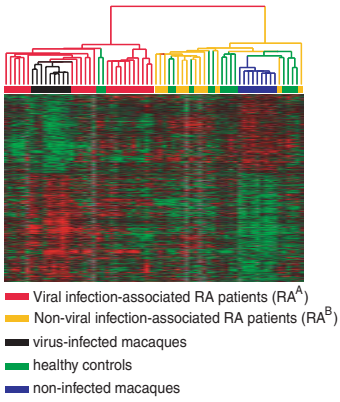


Figure 1.
Subgroup of RA patients clusters together with smallpox-virus infected macaques. Two-way hierarchical clustering was performed on genes that are significantly changed in expression after smallpox infection of macaques (18). Data from the smallpox study were combined with expression data from RA patients and healthy controls (see Material and Methods).

A direct comparison between the RA^A and RA^B subgroup by SAM revealed a prominent downregulation of genes in the RA^A group (2003 in total) representing processes involved in nucleotide metabolism, including pre-mRNA processing, mRNA transcription and mRNA splicing (analysed in the PANTHER database, $p= 5 \times 10^{-24}$, data not shown). In the smallpox infected macaques nucleotide metabolism was also highly significant downregulated ($p=1.3 \times 10^{-11}$, data not shown).

In our further analysis we compared the different RA subgroups to the healthy controls to obtain a view of the biological aberrations compared to a non-disease state. To identify genes that show an abnormal expression in the RA^A subgroup we performed a statistical analysis by SAM, which revealed that 663

genes were upregulated in RA[A] patients, while 1246 genes were downregulated compared to the healthy controls. To interpret the biological significance of the list of differentially expressed genes we performed a Gene Ontology analysis in the PANTHER database (Table 2).

Table 2. Gene Ontology of differentially expressed genes compared to controls.

	CPR*	All patients	RAA	RAB
Upregulated Biological Processes				
Immunity and defense	x	1.61E-11	1.32E-13	1.52E-03
Interferon-mediated immunity	x	1.09E-06	3.54E-04	ns
Ligand-mediated signaling	x	3.52E-03	5.38E-04	ns
Cytokine/chemokine mediated immunity	x	1.63E-03	3.50E-03	ns
Macrophage-mediated immunity	x	6.83E-05	1.13E-02	1.17E-02
Transport		4.82E-04	1.29E-02	4.87E-02
Lipid and fatty acid transport		1.54E-04	4.85E-02	ns
Signal transduction	x	ns	4.44E-06	4.24E-03
Cell proliferation and differentiation	x	ns	8.97E-06	ns
Developmental processes	x	ns	5.35E-05	ns
Blood clotting		ns	1.06E-04	ns
Cell communication	x	ns	2.06E-03	4.32E-02
Intracellular signaling cascade	x	ns	5.97E-03	ns
Inhibition of apoptosis	x	ns	6.35E-03	ns
Cell adhesion		ns	8.08E-03	1.18E-03
Cell structure and motility		ns	8.68E-03	ns
Cell surface receptor mediated signal transduction	x	ns	1.05E-02	ns
Mesoderm development	x	ns	1.69E-02	ns
Cell cycle		ns	1.81E-02	ns
Cell cycle control	x	ns	3.11E-02	ns
Apoptosis	x	ns	3.71E-02	ns
Downregulated Biological Processes				
Nucleoside, nucleotide and nucleic acid metabolism	x	ns	4.75E-09	ns
mRNA transcription	x	ns	1.16E-04	ns
mRNA transcription regulation	x	ns	3.66E-03	ns
Pre-mRNA processing		ns	5.71E-03	ns
Cell cycle		ns	1.91E-02	ns
Apoptosis	x	ns	4.33E-02	ns

Overview of the biological processes represented by the genes that were significantly different expressed in RA patient subgroups (Figure 1), compared to healthy individuals. Processes were identified by Panther Gene Ontology analysis. Processes with $P<0.05$ (after a Bonferroni correction) were selected as significant. ns: not significant. * Biological processes that were also represented by the Common Pathogen Response (CPR) (14) are indicated (x).

RA^{A} patients showed evidence of 21 upregulated processes, including Immunity and defense, Interferon-mediated immunity, Cytokine/chemokine mediated immunity, Cell proliferation and differentiation, (anti-) Apoptotic processes and Blood clotting. Among the 6 downregulated processes were several processes that indicated reduced transcriptional activity, cell cycling and apoptotic activity (in contrast to the upregulated genes these were mainly genes with pro-apoptotic activity). Within the list of significantly downregulated genes in the RA^{A} patient group were several genes suggesting a reduced cytotoxic response mediated by γδ T cells, including CCL5/RANTES, TARP/TCRgammaV9, perforin, NKG2D/KLRK1, and granzyme B, K and M.
The same analysis for the RA^{B} patients revealed only 375 upregulated genes and 179 downregulated genes compared to healthy controls (Table 3). Accordingly, there were fewer processes that discriminated RA^{B} patients from controls (increased Cell adhesion, Cell communication, Signal transduction, and Macrophage-mediated immunity) (Table 2).

Table 3. Overview of statistical analysis by SAM of the different RA subgroups (as defined in Figure 1), compared to healthy individuals.

	RA N=35	RA^{A} N=23	RA^{B} N=12	IFN^{high} N=20
No. Upregulated genes	259	663	375	484
No. Downregulated genes	318	1246	179	229
Total number of genes	577	1909	554	713

Our previous analysis on the IFN^{high} RA group (15) is included for comparison. The number of genes is indicated after a two-class unpaired SAM analysis using a false discovery rate of 5%.

To substantiate these findings we also performed gene ontology analysis on common pathogen response genes, i.e. a set of genes that is commonly induced by viral and bacterial triggers through activation of different TLRs (14). In comparison with the processes represented by the RA patients, we observed that in the RA^{A} group, 18 of the identified processes were shared with the common pathogen response, while in the RA^{B} group only 4 processes were identified in the common pathogen response (Table 2).

Finally, to gain more insight in the factors that could have induced the increased expression of genes in the RA^{A} patients we performed an analysis on transcription factor binding sites (TFBS) that are overrepresented in the genes, which were upregulated in the RA^{A} patients compared to the healthy individuals (by rVISTA (22) at http://genome.lbl.gov/vista/index.shtml, data not shown). The most significantly enriched regulatory elements in the 500 bp upstream regions of the overexpressed genes were ICSBP/IRF8, IRF-1, ISRE and NFκB binding sites. This indicates that agents that activate the NFκB-pathway and induce IFN type I production, are the most

important contributors to the expression signature of RA^A patients. Because these pathways are typical for TLR activation, we next performed this analysis on the common pathogen response gene set (14), The search for TFBS overrepresentation in the promoter regions of the common pathogen response genes revealed an almost identical set of most significant regulatory regions, namely NFκB, NFκB65, IRF-1 and ICSBP/IRF8 binding sites. The combined results suggest that activation of several TLRs, which mediate IFN type I induction and NFκB activation signals, such as TLR3, TLR4, TLR7 and TLR9, could potentially be responsible for the altered transcriptome in RA^A patients (23).

The most significantly over-represented TFBS in genes with a low expression in RAA patients were the binding sites for E2F1DP1 and E2F1DP2, which are involved in regulation of the cell cycle, DNA synthesis and apoptosis. The reduced activity of genes harboring these TFBS corresponds with the reduced activity of the pathways in RAA patients identified by gene ontology analysis.

Relation to the IFN-type high and low RA subgroups

We previously performed an RA subclassification based on the expression of IFN-induced genes (15). To obtain an insight in the differences between IFN- and viral-determined subclassifcation, we calculated the average expression of 43 correlated IFN-induced genes, which we described previously (15) for the different RA subgroups (Figure 2A). As anticipated the RA^A group showed a significantly higher expression of the IFN-response genes, when compared to both the RA^B group and the controls, corresponding with the Gene Ontology analysis. Although the RA^A group as a whole displayed an enhanced IFN type I response, six of the IFN^{low} patients were also included in the RA^A group, while four of the IFN^{high} patients were not represented in the RA^A group. This indicates that the subclassification of RA patients into RA^A and RA^B is not solely based on the expression of IFN-induced genes, but that the entire pathogen response program contributes to the subclassification. Statistical comparison of all groups to the healthy individuals indicated that the RA^A group was the most different from healthy controls (Table 3).

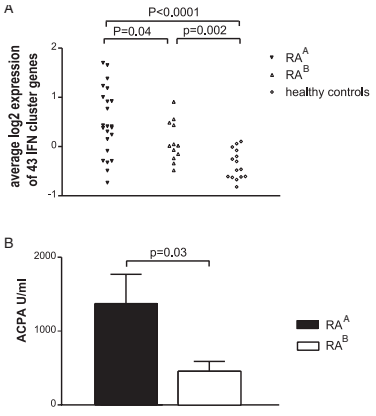


Figure 2.

A. IFN-response gene expression is enhanced in RA patients that clustered together with infected macaques (RA^A). The average expression ratio's of 43 IFN-response genes, that were identified in RA patients previously (15), are shown for the two patient groups and controls. B. ACPA levels are significantly higher in the RA patients that clustered together with infected macaques (RA^A). Data indicate the levels of antibodies to citrullinated proteins measured in the sera of RA patients

Clinical relevance of the patient stratification.
Next, the question arose whether the molecular stratification of RA was associated with clinical differences. To address this possibility, the relationship between the RA subtypes and clinical parameters, summarized in Table 1, was determined. Most strikingly, this analysis revealed that the two groups differed in ACPA levels. The RA^A group showed higher levels of ACPA ($p = 0.03$, Figure 2B) than the RA^B group.

Discussion

The combined cluster analysis of RA patients with smallpox virus infected macaques revealed a striking heterogeneity between RA patients indicated by the co-clustering of a subset of patients (RAA subgroup) with the virus-infected primates. The combined analysis of different datasets has already demonstrated its significance by the discovery of a common fibroblast serum response signature that is represented in tumor samples and predicts cancer progression (24). Our current analysis further indicates that combining different datasets powers disease subclassification and leads to the identification of an immune defense gene expression profile in peripheral blood cells of a subgroup of RA patients.
Infectious agents have long been considered as possible triggers for autoimmune responses (25-27). The higher load of EBV DNA identified in peripheral blood cells from RA patients (28) would support this suggestion. Olsen et al proposed a "viral signature" as a typical feature of early RA, based on the observation that the transcriptome of peripheral blood cells from early RA patients partly overlaps with genes induced by Influenza (29) . Although our study was not designed to compare early with longstanding RA, the disease duration was shorter in the RAA compared to the RAB group (57 vs 132 months), but did not reach statistical significance ($p=0.099$). Further support for an infectious agent is delivered by the remarkable resemblance of the gene expression signature of the RAA group with the transcriptional program of the common pathogen response (14). In addition, the most important transcription factor binding sites in the upstream regions of the genes expressed by the RAA patients correspond with the regulatory regions used by genes that comprise the common pathogen response; the binding sites for ICSB/IRF8, IRF1, and NFκB suggests the involvement of TLR activation in the RAA group. Interestingly, in genes with a low expression in RAA patients the most significant transcription factor binding sites were E2F1DP1 and E2F1DP2. These factors are normally able to inhibit the NFκB survival signal (30). A reduced activity of these transcription factors would imply that an uninhibited survival signal could be delivered by NFκB in the RAA patients.
Another resemblance with a viral infection is the observation that the RAA patient group showed signs of a reduced cytotoxic response mediated by γδ T cells, suggested by the reduced expression of CCL5/RANTES, TARP/TCRgammaV9, perforin, NKG2D/KLRK1, and granzyme B, K and M. This finding is in line with the

previously identified low levels of γδ T cells in peripheral blood of RA patients (31) Furthermore, a lack of expression of granzyme A, RANTES/CCL5, and TCRB, during hepatitis C virus infection in non-human primates has shown to be associated with the inability to clear the infection (32). Therefore, RAA patients may be more prone to viral persistence because of the lack of effective virus clearance. In support of this, a reduced anti-viral response, indicated by the decreased T cell precursor frequencies to EBV gp110 in established RA has been reported (33). Although we were not able to directly compare our data to a chronic viral infection, several characteristics of acute viral infections have also been observed in chronic viral infections. For instance, patients with chronic hepatitis C virus infection show reduced numbers of circulating γδ T cells as well (34). In addition, the type I IFN-response expression profile, is also convincingly present in PBMC from patients with a chronic hepatitis C infection (35)

In summary, the RAA patients show several characteristics of viral infections, a) the increased type I IFN signature, b) a reduced / gene expression signature, c) reduced transcript levels of HLA class II molecules (data not shown), and d) reduced transcriptional activity(34;36;37).

The fact that the RAB patients were not very different from controls in immunity-related transcript levels, may suggest that local inflammation in the joints is less severe, and therefore not as profoundly reflected in peripheral blood cells.

In these patients a more aggressive type of fibroblast-like synoviocytes that are relatively immune-independent could possibly drive the disease, as we suggested earlier (38)

Although our current analysis supports the suggestion that an infectious agent is involved in the development of arthritis, the study design does not allow for a firm conclusion to be drawn. Formally we cannot exclude the possibility that the pathogen-response signature may have been triggered by endogenous danger signals that are released during inflammation, independent of infectious agents, such as heat shock proteins and DNA or RNA from damaged or dying cells (39). Recently, defective degradation of endogenous DNA from apoptotic cells has been identified as a trigger for arthritis and the generation ACPA in an animal model, leading to increased IFN type I and TNFα release by an unknown mechanism (40). We found evidence for diminished apoptosis in the RAA group, which suggests that a defect in apoptosis may contribute to the auto-immune process, as has been suggested before (41). The increased ACPA levels in the RAA patients might therefore be related to defects in the apoptotic pathway. Of interest is the recent observation that approximately half of the RA patients show antibody reactivity to citrullinated Epstein-Barr virus (EBV)-derived peptide, which correlates with ACPA levels (42). These findings might suggest that a viral agent could also play a role in the induction of ACPA antibodies.

We are now facing the challenge of unraveling the etiology of the pathogen-response signature in RA.

Funding

This study was supported in part by the EURO-RA Marie Curie Trainings network, the European Community's FP6 funding (Autocure), the Netherlands Organization for Health Research and Development (ZonMw) in assignment of the Netherlands Organization for Scientific Research (NWO, grant number 945-02-029), the Innovation Oriented research Program (IOP) on Genomics and Centre for Medical Systems Biology (a center of excellence approved by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research). This publication reflects only the authors' views. The European Community is not liable for any use that may be made of the information herein.

Acknowledgements

We would like to thank the following authors for making their microarray data publicly available: RG Jenner (Cambridge, USA) and KH Rubins (Stanford, USA)

Reference List

(1) Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM et al. Lymphoid neogenesis in rheumatoid synovitis. J Immunol 2001; 167(2):1072-1080.

(2) Timmer TC, Baltus B, Vondenhoff M, Huizinga TW, Tak PP, Verweij CL et al. Inflammation and ectopic lymphoid structures in rheumatoid arthritis synovial tissues dissected by genomics technology: identification of the interleukin-7 signaling pathway in tissues with lymphoid neogenesis. Arthritis Rheum 2007; 56(8):2492-2502.

(3) Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group 1. **N Engl J Med 2000; 343(22):1594-1602.**

(4) Kremer JM, Westhovens R, Leon M, Di Giorgio E, Alten R, Steinfeld S et al. **Treatment of rheumatoid arthritis by selective** inhibition of T-cell activation with fusion protein CTLA4Ig. N Engl J Med 2003; 349(20):1907-1915.

(5) Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. N Engl J Med 2004; 350(25):2572-2581.

(6) Zendman AJ, van Venrooij WJ, Pruijn GJ. Use and significance of anti-CCP autoantibodies in rheumatoid arthritis. Rheumatology (Oxford) 2006; 45(1):20-25.

(7) Hueber W, Kidd BA, Tomooka BH, Lee BJ, Bruce B, Fries JF et al. Antigen microarray profiling of autoantibodies in rheumatoid arthritis. Arthritis Rheum 2005; 52(9):2645-2655.

(8) Firestein GS, Zvaifler NJ. How important are T cells in chronic rheumatoid synovitis?: II. T cell-independent mechanisms from beginning to end. Arthritis Rheum 2002; 46(2):298-308.

(9) Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 1993; 32(10):903-907.

(10) Deighton CM, Wentzel J, Cavanagh G, Roberts DF, Walker DJ. Contribution of inherited factors to rheumatoid arthritis. Ann Rheum Dis 1992; 51(2):182-185.

(11) Ramirez AS, Rosas A, Hernandez-Beriain JA, Orengo JC, Saavedra P, de la FC et al. Relationship between rheumatoid arthritis and Mycoplasma pneumoniae: a case-control study. Rheumatology (Oxford) 2005; 44(7):912-914.

(12) Haier J, Nasralla M, Franco AR, Nicolson GL. Detection of mycoplasmal infections in blood of patients with rheumatoid

arthritis. Rheumatology (Oxford) 1999; 38(6):504-509.

(13) Mehraein Y, Lennerz C, Ehlhardt S, Remberger K, Ojak A, Zang KD. Latent Epstein-Barr virus (EBV) infection and cytomegalovirus (CMV) infection in synovial tissue of autoimmune chronic arthritis determined by RNA- and DNA-in situ hybridization. Mod Pathol 2004; 17(7):781-789.

(14) Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. Nat Rev Microbiol 2005; 3(4):281-294.

(15) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. **Ann Rheum Dis 2007; 66(8):1008-1014.**

(16) Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D et al. **The Stanford Microarray Database:** implementation of new analysis tools and open source release of software. Nucleic Acids Res 2007; 35(Database issue): D766-D770.

(17) Nielsen TO, West RB, Linn SC, Alter O, Knowling MA, O'Connell JX et al. Molecular characterisation of soft tissue tumours: a gene expression study. Lancet 2002; 359(9314):1301-1307.

(18) Rubins KH, Hensley LE, Jahrling PB, Whitney AR, Geisbert TW, Huggins JW et al. The host response to smallpox: analysis of the gene expression program in peripheral blood cells in a nonhuman primate model. Proc Natl Acad Sci U S A 2004; %19;101(42):15190-15195.

(19) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-14868.

(20) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-5121.

(21) Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 2005; 33(Database issue):D284-D288.

(22) Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM. rVista for comparative sequence-based discovery of functional transcription factor binding sites. Genome Res 2002; 12(5):832-839.

(23) Moynagh PN. TLR signalling and activation of IRFs: revisiting old friends from the NF-kappaB pathway. Trends Immunol 2005; 26(9):469-476.

(24) Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K et al. Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds. PLoS Biol 2004; 2(2):E7.

(25) Wucherpfennig KW. Mechanisms for the induction of autoimmunity by infectious agents. J Clin Invest 2001; 108(8):1097-1104.

(26) Cacoub P, Poynard T, Ghillani P, Charlotte F, Olivi M, Piette JC et al. Extrahepatic manifestations of chronic hepatitis C. MULTIVIRC Group. Multidepartment Virus C. Arthritis Rheum 1999; 42(10):2204-2212.

(27) Stahl HD, Hubner B, Seidl B, Liebert UG, van dH, I, Wilbrink B et al. Detection of multiple viral DNA species in synovial tissue and fluid of patients with early arthritis. Ann Rheum Dis 2000; 59(5):342-346.

(28) Balandraud N, Roudier J, Roudier C. Epstein-Barr virus and rheumatoid arthritis. Autoimmun Rev 2004; 3(5):362-367.

(29) Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J et al. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. Ann Rheum Dis 2004; 63(11):1387-1392.

(30) Hitchens MR, Robbins PD. The role of the transcription factor DP in apoptosis. Apoptosis 2003; 8(5):461-468.

(31) Lunardi C, Marguerie C, Walport MJ, So AK. T gamma delta cells and their subsets in blood and synovial fluid from patients with rheumatoid arthritis. Br J Rheumatol 1992; 31(8):527-530.

(32) Su AI, Pezacki JP, Wodicka L, Brideau AD, Supekova L, Thimme R et al. Genomic analysis of the host response to hepatitis C virus infection. Proc Natl Acad Sci U S A 2002; 99(24):15669-15674.

(33) Toussirot E, Wendling D, Tiberghien P, Luka J, Roudier J. Decreased T cell precursor frequencies to Epstein-Barr virus glycoprotein Gp110 in peripheral blood correlate with disease activity and severity in patients with rheumatoid arthritis. Ann Rheum Dis 2000; 59(7):533-538.

(34) Par G, Rukavina D, Podack ER, Horanyi M, Szekeres-Bartho J, Hegedus G et al. Decrease in CD3-negative-CD8dim(+) and Vdelta2/Vgamma9 TcR+ peripheral blood lymphocyte counts, low perforin expression and the impairment of natural killer cell activity is associated with chronic hepatitis C virus infection. **J Hepatol 2002; 37(4):514-522.**

(35) Mihm S, Frese M, Meier V, Wietzke-Braun P, Scharf JG, Bartenschlager R et al. **Interferon type I gene expression in chronic** hepatitis C. Lab Invest 2004; 84(9):1148-1159.

(36) Lyles DS. Cytopathogenesis and inhibition of host gene expression by RNA viruses. Microbiol Mol Biol Rev 2000; 64(4):709-724.

(37) Poccia F, Boullier S, Lecoeur H, Cochet M, Poquet Y, Colizzi V et al. Peripheral V gamma 9/V delta 2 T cell deletion and anergy to nonpeptidic mycobacterial antigens in asymptomatic HIV-1-infected persons. J Immunol 1996; 157(1):449-461.

(38) Kasperkovitz PV, Timmer TC, Smeets TJ, Verbeet NL, Tak PP, van Baarsen LG et al. Fibroblast-like synoviocytes derived from patients with rheumatoid arthritis show the imprint of synovial tissue heterogeneity: evidence of a link between an increased myofibroblast-like phenotype and high-inflammation synovitis. Arthritis Rheum 2005; 52(2):430-441.

(39) Kol A, Bourcier T, Lichtman AH, Libby P. Chlamydial and human heat shock protein 60s activate human vascular endothelium, smooth muscle cells, and macrophages. J Clin Invest 1999; 103(4):571-577.

(40) Kawane K, Ohtani M, Miwa K, Kizawa T, Kanbara Y, Yoshioka Y et al. Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages. Nature 2006; 443(7114):998-1002.

(41) Maas K, Chan S, Parker J, Slater A, Moore J, Olsen N et al. Cutting edge: molecular portrait of human autoimmune disease. J Immunol 2002; 169(1):5-9.

(42) Anzilotti C, Merlini G, Pratesi F, Tommasi C, Chimenti D, Migliorini P. Antibodies to viral citrullinated peptide in rheumatoid arthritis. J Rheumatol 2006; 33(4):647-651.

.

Chapter 3.3

Synovial tissue heterogeneity in rheumatoid arthritis in relationship to disease activity and biomarkers in peripheral blood

3

Lisa G.M. van Baarsen[1,2], Carla A. Wijbrandts, MD[2], Trieneke C.G. Timmer[1], Tineke C.T.M. van der Pouw Kraan, PhD[1], Paul P. Tak, MD, PhD[2] and Cornelis L. Verweij, PhD[1]

[1]VU Medical Center, Amsterdam, Netherlands
[2]Academic Medical Center/University of Amsterdam, Amsterdam, The Netherlands

Submitted for publication

Abstract

Objective: We investigated the clinical relevance of synovial tissue subtypes in rheumatoid arthritis (RA) and searched for peripheral blood (PB) markers that may serve as biomarkers associated with synovial tissue subtypes.

Methods: Gene expression analysis using cDNA microarrays was applied on synovial tissue biopsies obtained by arthroscopy and paired PB samples derived from 17 RA patients. Molecular tissue subtypes were correlated with histological parameters, disease characteristics and PB markers.

Results: Cluster analysis based on gene expression profiles and independently on immunohistochemical data resulted in an identical subdivision of high and low inflammation tissues. The high inflammation tissue type was predominantly found in patients with a higher disease activity score and shorter disease duration. They also had increased levels of C-reactive protein (CRP); higher erythrocyte sedimentation rates (ESR) and increased numbers of platelets. At the single gene level the PB gene expression profiles yielded no statistically significant differences between the two patient groups. However, pathway-level analysis of the PB profiles revealed that tissue subtypes correlated with differential expression of sets of genes involved in protein biosynthesis and developmental processes.

Conclusion: High tissue inflammation is associated with more severe disease and shorter disease duration. Differential expression of sets of genes in PB is associated with specific subtypes of tissue inflammation, which cannot be detected at the single gene level in PB.

Introduction

Rheumatoid arthritis (RA) is a systemic inflammatory autoimmune disease of unknown origin, affecting the synovium in multiple joints. The heterogeneous nature of the disease is reflected by the variability in clinical presentation (1;2) and treatment response (3).

The synovial hypertrophy in RA is largely due to the increased recruitment of inflammatory cells that migrate from the blood into the synovium although the histopathology varies widely between patients (4-6). Accordingly, gene expression profiling in tissues from patients with RA revealed that tissue heterogeneity is reflected at the molecular level (7). Patients could be classified in at least two groups based on their gene expression profile in synovial tissue. One group of patients showed increased expression of inflammatory genes whereas the other group showed increased expression of genes involved in tissue remodeling (7). These results point towards the possibility of different pathogenic mechanisms playing a role in RA.

Synovial tissue heterogeneity reflects the severity of synovial and systemic inflammation (8;9) and could represent differences in the underlying pathology, requiring different treatments. The analysis of biomarkers in peripheral blood

(PB) rather than synovial tissue may facilitate studies on the relationship between pathogenetic subsets and clinical signs and symptoms in large patient groups, as synovial biopsy is not available in all centers. The notion that gene expression profiles in PB may reflect interindividual heterogeneity is supported by the presence of type I interferon (IFN)-induced genes in the PB of a subset of RA patients (10). However, the clinical relevance of a type I IFN-response in RA is at present unknown.
The aim of this study was to investigate the clinical relevance of synovial tissue subtypes in RA and to search for PB markers that may serve as biomarker for synovial tissue subtypes. Therefore, we applied gene expression analysis using cDNA microarrays on a unique set of paired synovial tissue biopsies and peripheral whole blood samples derived from 17 RA patients.

▪ Materials & Methods

Patients
Consecutive patients with RA according to the ACR criteria were enrolled in the study at the outpatient clinic of the Academic Medical Center (AMC) in Amsterdam over a period of one year. Inclusion criteria were: 18-85 years of age, a failure of at least two disease-modifying anti-rheumatic drugs (DMARDs) including methotrexate (MTX), and active disease (disease activity score in 28 joints (DAS28) ≥ 3.2). Patients with a history of an acute inflammatory joint disease of different origin or previous use of a TNF blocking agent were excluded. Patients were on stable maximally tolerable MTX treatment. Whole blood samples were obtained using PAXgene tubes (PreAnalytix, GmbH, Germany) from 17 RA patients. From the same patients synovial tissue was obtained by arthroscopic synovial biopsy. All patients gave written informed consent and the study protocol was approved by the Medical Ethics Committee (AMC).

Sample preparation and hybridization for microarray analysis
Total RNA from synovial tissue biopsies was isolated as described previously (11). In total 2.5 ml blood was drawn in PAXgene blood RNA isolation tubes (PreAnalytix, GmbH, Germany) and stored at -20°C. Prior to RNA isolation tubes were thawed for 2 hours at room temperature. Next, total RNA was isolated using the PAXgene RNA isolation kit according to the manufacturer's instructions including a DNAse (Qiagen, Venlo, Netherlands) step to remove genomic DNA. Quantity and purity of the RNA was tested using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, DE).
We used 43K cDNA microarrays from the Stanford Functional Genomics Facility (http://microarray.org/sfgf/) printed on aminosilane-coated slides containing ~20.000 unique genes. First DNA spots were UV-cross linked to the slide using 150-300 mJoules. Sample preparation and microarray hybridization was performed as described previously (12;13).

Microarray data analysis

Data storage and filtering was performed using the Stanford Microarray Database (14) (http://genome-www5.stanford.edu//) as described previously (15). Raw data can be downloaded from the publicly accessible Stanford database website. We used the Q-score tool from the database as a quality measure to remove low quality spots. Q-score determined the appropriate filter criteria for: the regression correlation between channels 1 and 2, the background settings and the minimal channel intensities. After removing the low quality spots, data values with the same Unigene Identifier were averaged and all array data was median centered (genes and arrays). Statistical Analysis of Microarrays (16) (SAM) was used to determine significantly differential expressed genes. A gene was considered as significantly differentially expressed if the False Discovery Rate (FDR) was equal to or less than 5%. Cluster analysis (17) was used for subclassification of patient groups and to define clusters of coordinately changed genes after which the data was visualized using Treeview. To interpret our results, we applied PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (Applied Biosystems, Foster City, CA) at http://PANTHER.appliedbiosystem.com (18;19). This analysis uses the binomial statistics tool to compare the list of up- or down-regulated genes to a reference list to statistically determine over- or under- representation of PANTHER classification categories such as biological processes. A Bonferroni correction was applied to correct for multiple comparisons and after classification a significant p-value ($p<0.05$) indicates that a given category may be of biological interest.

Synovial biopsy and immunohistochemistry

The arthroscopy, tissue sampling and storage were performed as previously described (20). The cellular infiltrate and the presence of adhesion molecules were determined by immunohistochemical analysis as described before (21). The following antibodies were used to analyze the cellular infiltrate: anti-CD68 (EBM11: DAKO, Denmark) to detect macrophages, anti-CD3 (SK7, Becton Dickinson (BD), CA) for T cells, anti-CD22 (CLB-B-ly/1,6B11, the Netherlands) for B cells, anti-CD38 (HB7, BD) for plasma cells, and anti-CD163 (Ber-MAC3; DAKO) for resident tissue macrophages. For the detection of cytokines and adhesion molecules we used: anti-human TNFα (52B83; Monosan, Belgium), anti-ICAM-1 (MEM111, Sanbio, Belgium), anti-VCAM (1G11B1, Sanbio) and anti-E-selectin (BBIG-E4, R&D, Minneapolis, MN). Staining of cellular markers was performed using a three-step immunoperoxidase method. For staining of cytokines and adhesion molecules biotinylated tyramine was used for amplification. For control sections the primary antibody was omitted or irrelevant immunoglobulins were applied. Images were acquired and analyzed by computer-assisted image analysis using a Syndia algorithm on a Qwin-based analysis system (Leica, Cambridge, UK) (22). Cellular markers were expressed as positive cells/mm^2 (counts/mm^2). Staining of cytokines and adhesion molecules was expressed as integrated optical density/mm^2 (IOD/mm^2). CD68+ macrophages

and TNFα expression were analyzed separately in the intimal lining layer and the synovial sublining.
The presence of germinal centers in the synovium was assessed by specific staining for follicular dendritic cells (FDCs) using anti-CD21L (kindly provided by Dr. Y-J. Liu) as previously described (23).

Real-time PCR
In one 48-wells PCR plate (Greiner Bio-One GmbH, Frickenhausen, Germany) 0.5 μg of total RNA was reverse transcribed into cDNA using the Revertaid H-minus cDNA synthesis kit (MBI Fermentas, St. Leon-Rot, Germany) according to the manufacturer's instructions. Quantitative real-time PCR was performed using an ABI Prism 7900HT Sequence detection system (Applied Biosystems) using SybrGreen (Applied Biosystems). Primers were designed using Primer Express software and guidelines (Applied Biosystems). To calculate arbitrary values of mRNA levels and to correct for differences in primer efficiencies a standard curve was constructed. Expression levels of target genes were expressed relative to GAPDH.

MMP3 ELISA
Serum levels of MMP3 were measured using a commercially available MMP3 cytoset (Invitrogen, Breda, Netherlands), which detects both the inactive (pro-MMP3) and the active form of MMP3. The ELISA was performed according to the manufacturer's instructions using ten times diluted serum samples and duplicate measurements.

Statistical analysis
Data were analyzed using the software programs Graphpad Prism 4 (Graphpad Software, Inc., La Jolla, CA) and SPSS version 14.0 (Chicago, IL). First, the data were checked for normal distribution. Next, two-group comparison of normally distributed data was analyzed using the non-parametric independent sample t-test including Levene's test for equality of variances. Not normally distributed data were analyzed using the two-independent-samples Mann-Whitney U test. Correlation analyses were performed using Spearman or Pearson correlation analysis as appropriate. Differences were considered statistically significant with P-values less than 0.05.

▪ Results

Molecular profile of synovial tissue
In order to create a framework for an explorative approach in search for PB biomarkers that correlate with synovial tissue subtypes in RA, we studied molecular and biological features of paired PB and synovial tissue samples. Gene expression profiles were analyzed using cDNA microarrays from synovial biopsies

of 17 patients with RA obtained by arthroscopy. To study tissue heterogeneity we selected genes, which gene expression levels varied at least two-fold from the median expression value in at least three patients. The remaining 256 genes were unsupervised clustered using two-way hierarchical cluster analysis and visualized by Treeview (Figure 1). In accordance with previous findings (7), this molecular approach revealed the presence of at least two subgroups (I and II) of patients based on their differential tissue gene expression profile.

To provide insight into the significantly differentially expressed genes between the two identified tissue groups, we used SAM: 538 genes were significantly higher expressed in group I, whereas 1009 genes were significantly higher expressed in tissue group II. Genes previously associated with the pathogenesis of RA and expressed at higher levels in tissue group I included CILP and COMP. Conversely, genes expressed at higher levels in tissue group II and associated with RA were e.g. MMP3, SPP1 (also known as Osteopontin), FCGR1A and several genes encoding for immunoglobulins.

To interpret the biological function of these significantly differentially expressed genes we applied PANTHER's classification system. This pathway-level analysis revealed that the 538 genes that were expressed at significantly higher levels in tissue group I displayed an overrepresentation of genes involved in among others; developmental processes and neurogenesis. The 1009 genes higher expressed in tissue group II could be classified into several inflammation related processes such as immunity and defense, T-cell and B-cell mediated immunity, IFN mediated immunity and others (Table 1).

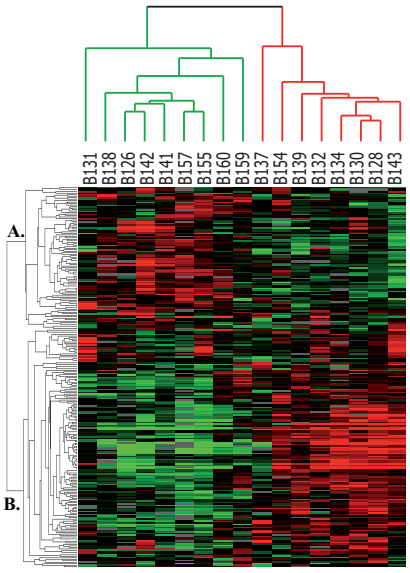


Figure 1. Molecular heterogeneity of RA synovium
Two-way hierarchical cluster analysis was used to subclassify patients based on correlated expression profiles in synovial tissues. Each column represents the data of one array/sample and each row shows the relative expression levels of a single gene for all samples. Red color means a relatively higher expression, green color stands for a relatively lower expression and a black color black indicates that the expression level is equal to the median expression level across all samples. Grey indicates missing values. Genes were selected if expression levels deviated from the median expression level at least twofold in at least three tissues. The resulting 256 genes were clustered and this resulted in a subclassification of patients based on the differential expression of genes present in gene clusters A and B.

Table 1. PANTHER analysis

Biological process enriched in tissue group I	NCBI: H.sapiens genes (REF)	In SAM list	expected	p-value
Developmental processes	2152	76	32	6.2E-11
Ectoderm development	692	27	10	0.00115
Signal transduction	3406	79	51	0.00139
Mesoderm development	551	23	8	0.00193
Neurogenesis	587	24	9	0.00239
Cell structure and motility	1148	34	17	0.00449
Cell motility	352	17	5	0.00449
Cell communication	1213	37	18	0.00583
Lipid, fatty acid and steroid metabolism	770	25	12	0.00924
Cell structure	687	24	10	0.0198
Cell adhesion	622	20	9	0.0408

Biological process enriched in tissue group II	NCBI: H.sapiens genes (REF)	In SAM list	expected	p-value
Immunity and defense	1318	159	44	1.06E-43
T-cell mediated immunity	194	44	6	1.41E-20
Signal transduction	3406	194	113	5.48E-13
Cytokine and chemokine mediated signaling pathway	252	35	8	6.44E-10
B-cell- and antibody-mediated immunity	97	22	3	6.62E-10
Interferon-mediated immunity	63	17	2	1.30E-08
Natural killer cell mediated immunity	74	18	2	1.90E-08
Macrophage-mediated immunity	140	24	5	2.07E-08
MHCII-mediated immunity	34	11	1	6.05E-06
Cell surface receptor mediated signal transduction	1638	96	54	8.14E-06
Cell communication	1213	72	40	2.69E-04
Ligand-mediated signaling	421	33	14	1.49E-03
Apoptosis	531	36	18	1.86E-03
MHCI-mediated immunity	22	7	1	2.19E-03
Intracellular signaling cascade	871	53	29	3.33E-03

Comparative analysis between gene expression profiling and immuno-histochemical analysis

To perform a histological validation of the molecular classification of the two tissue subtypes we performed immunohistochemical analysis for the cellular markers CD163 and CD68 (macrophages), CD22 (B-cells), CD3 (T-cells) and CD38 (plasma cells). In addition, the tissues were stained for E-Selectin, TNF, VCAM and ICAM. Two-way hierarchical cluster analysis of the immunohistochemical scores revealed a clear subclassification of the tissues in at least two groups, which was mostly

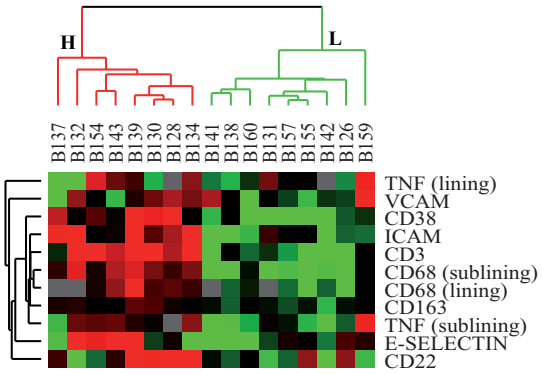


Figure 2. Subclassification based on immunohistochemistry

Synovial tissues were stained for CD163, CD22, CD3, CD38, CD68, TNF, E-Selectin, VCAM and ICAM. The amount of staining was quantified as number of positive cells per mm^2 or IOD per mm^2. In order to transform these values into comparable numbers for clustering, for each parameter the mean value was calculated after which the value per patient was divided by the mean and log base two transformed. These numbers were two-way hierarchical clustered and visualized by Treeview.

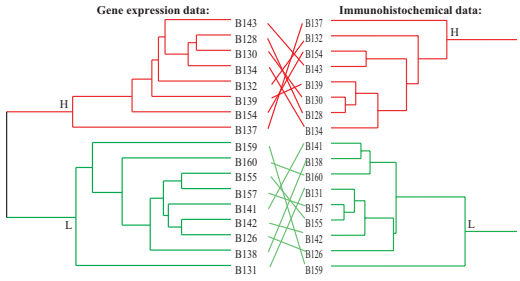


Figure 3. Molecular subclassification fits with differences detected at protein level

The left side of the figure shows the subclassification based on differences in gene expression profiles (Figure 1). The right side shows subclassification based on differences detected at protein level by immunohistochemistry (Figure 2). Patients were similarly divided into two groups (high and low inflammatory) based on differences detected at transcript and protein level.

based on differences in the amount of infiltrating immune cells and the expression of E-selectin (Figure 2).

Comparative analysis of subclassification into two groups between gene expression profiling and immunohistochemical analysis revealed assignment of identical subgroups (Figure 3). This means that gene expression profiling provides a molecular signature of high inflammatory and low inflammatory synovium.

Knowing that the organization of cellular infiltrates in the synovium may differ between RA patients (8;9;24;25), we next investigated the relationship between the high and low inflammatory subtypes and the level of T and B cell organization observed in tissue. As anticipated, all tissues from the low inflammatory group displayed a diffuse type of cellular infiltrate. All but one of the high inflammatory tissues displayed either an aggregated or germinal center-like organization. Seven high inflammatory tissues contained aggregated clusters of T and B cells and two tissues stained positive for CD21L which is selectively present on FDCs suggestive for the presence of germinal centers in those two high inflammatory tissues (B130 and B134, data not shown). Only one high inflammatory tissue (B137) displayed a diffuse tissue type, but this sample already clustered apart from the other high inflammatory group already reflecting the difference. Collectively, the

subclassification correlated with the cellular infiltrate in the synovial tissues: the low inflammatory tissues were characterized by a diffuse organization of infiltrating cells and the high inflammatory tissue by an aggregated and germinal-center like organization.

Different tissue types in relationship to the clinical phenotype

Next the clinical characteristics between patients with high and low inflammatory tissue subtypes were compared (Table 2). Patients with a high inflammatory tissue type had a higher DAS28, a higher number of tender and swollen joints, higher visual analog scale for disease activity (VAS), as well as increased levels of ESR, CRP, and platelets. In addition, these patients had a shorter disease duration than the low inflammatory patient group. Autoantibody levels were not significantly different between the two groups, although there was a trend towards higher levels in the high inflammation tissue group.

Table 2. Clinical and demographic differences between tissue types

	High or low inflammatory tissue	
Clinical characteristics	low	high
Age (years)	55 (50-60)	52 (35-56)
DAS28	4.6 (4.5-5.6)	7.2 (6.2-7.3)*
Disease duration (months)	126 (48-163)	42 (21-60)*
Methotrexate (mg/week)	20 (12.5-30)	25 (18.8-25)
IgM rheumatoid factor (U/L)	28 (14-133)	195 (108-405)
ACPA (U/L)	342 (48-595)	736 (40-1211)
Tender joint count	7 (5-12)	16 (12-16)*
Swollen joint count	11 (6-15)	16 (13-17)
VAS (mm)	34 (29-40)	71 (63-79)*
HAQ-DI (0-3)	0.75 (0.5-1.5)	1.4 (1.1-1.9)
Morning stiffness (minutes)	30 (10-30)	45 (14-150)
ESR (mm/ hour)	23 (20-26)	44 (35-62)*
CRP (mg/L)	6 (5-11)	40 (14-47)*
Leukocytes (10^9/L)	6.5 (5.9-7.1)	9.7 (6.2-11.2)
Neutrophils (%)	71 (68-74)	74 (62-78)
Lymphocytes (%)	20 (16-25)	18 (13-26)
Monocytes (%)	6 (6-7)	6 (6-8)
Thrombocytes	234 (233-292)	373 (350-412)***

Values are presented as median and interquartile range (IQR; 25^{th}-75^{th} percentile). DAS28=disease activity score in 28 joints, ACPA=antibodies against cyclic citrullinated proteins, VAS=visual analog scale, HAQ-DI= health assessment questionnaire-disability scores, ESR=erythrocyte sedimentation rate, CRP=C-reactive protein;

* means significantly different between low and high inflammatory group P-value < 0.05, *** P-value < 0.001

Peripheral blood markers reflecting tissue inflammation state

In search for novel PB biomarkers, we analyzed the paired PB gene expression profiles in a supervised manner, i.e. driven by the tissue classification. Comparing

the blood profiles of the high inflammatory patients with those of the low inflammatory patients revealed no individual significantly differentially expressed genes. These findings indicate that differential tissue pathology is not reflected in the PB by differential expression of single genes.

Subsequently, we applied PANTHER pathway-level analysis to search for significantly up- or downregulated biological processes in the blood compartment that could discriminate between the tissue subtypes. This analysis showed that the PB profile of the high inflammatory tissue patient group displays a significant overrepresentation of genes involved in protein biosynthesis ($P < 0.0001$). Conversely, the PB profiles of the low inflammatory tissue group exhibited an enriched expression of genes involved in developmental processes ($P = 0.055$). Thus, whereas the molecular differences in PB cells between high and low inflammatory tissues are not reflected by significant differential expression of individual genes, pathway level-analysis did provide evidence for a significant association with distinct biological processes.

In addition we searched for serum proteins which could be differentially expressed between the high and low inflammatory tissue patient groups. Therefore, we measured serum levels of MMP3, which was one of the key genes that discriminated between the low and high inflammatory tissue groups. The significant differences in transcript levels for MMP3 measured by microarray analysis were confirmed by quantitative realtime PCR (qPCR, Figure 4A). Although serum levels of MMP3 protein correlated with the transcript levels produced in the synovium (Pearson's $R = 0.61$, $P = 0.015$), no statistically significant difference was observed comparing serum samples from patients with the two different tissue types (Figure 4B).

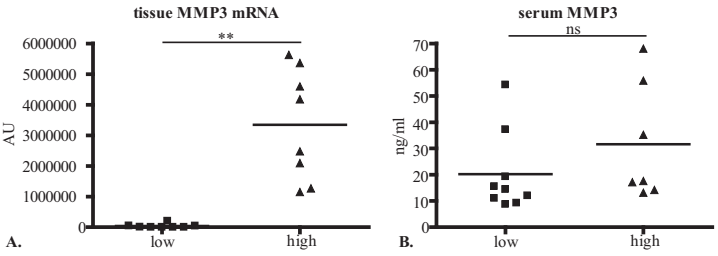


Figure 4. MMP3 levels

Patients were divided into high and low inflammatory tissue groups based on synovial tissue analysis and MMP3 levels were compared. Transcript levels of MMP3 measured by qPCR (A). Serum levels of MMP3 measured by ELISA (B). ** $p<0.01$ unpaired t test

Discussion

Gene expression profiling of synovial tissues biopsies confirmed and further extended our previous findings (7;11) of two different synovial tissue types in RA: high and low inflammatory tissues. The subclassification based on molecular markers exactly matched with the subclassification based on immunohistochemical

markers for cell infiltration. Accordingly, gene expression profiling provides a molecular signature of high inflammatory versus low inflammatory synovium. High inflammatory tissues were associated with a higher disease activity score and shorter disease duration. In addition, the high inflammatory tissue types correlated with increased levels CRP, ESR, and platelets. However, at the single gene expression level, no significant differences in paired PB blood gene expression profiles were observed between the two tissue types. Pathway-level analysis revealed that tissue subtypes correlated with differential expression of sets of genes involved in protein biosynthesis and developmental processes.

These results confirm the existence of two different pathogenic mechanisms of disease, one inflammation-mediated and one immune-independent, as suggested previously (26;27). Collectively, the high inflammatory tissues displayed an upregulation of genes involved in several inflammation mediated processes such as (MHC-mediated) immunity and defense, T/B cell-mediated immunity, cytokine/chemokine-mediated signaling, IFN-mediated immunity, macrophage-mediated immunity and others. Compared to the high inflammatory tissues the low inflammatory tissues exhibit increased expression of genes involved in developmental processes and neurogenesis. The low inflammatory tissue type represents a pathogenic process mainly driven by deregulated resident cells such as fibroblasts leading to tissue destruction without tissue inflammation combined with lower disease activity, CRP and ESR. In line with these findings, it has been suggested that COMP, found here to be associated with low inflammatory tissue, is arthritogenic by itself in experimental models of arthritis (28). Moreover, patients with increased serum levels of COMP had a lower response to anti-TNF therapy, independent of CRP levels (29). These findings indicate that COMP may have a pathogenic role in RA independent of the underlying immune-mediated inflammatory processes.

We cannot exclude the possibility that the differences reflect two different phases of the disease. The observation that the low inflammatory tissues are associated with longer disease duration supports this notion. Temporal differences have been proposed previously (30;31): progression could occur from a T and B cell driven disease to a more immune-independent process that is mainly driven by autonomous fibroblast-like synoviocytes. We have previously shown that so-called early RA represents chronic synovitis; the features of synovial inflammation are on average similar between RA < 1 year's duration compared to RA > 5 years' duration (5;32). The results presented here suggest that the transition from immune-mediated inflammation to a more autonomous, immune-independent process may occur several years after onset of clinical signs and symptoms.

Several groups studied the relationship between MMP3 and inflammatory tissue injury (33), and MMP3 appears to play an important role in the pathogenesis of RA (34-36). Although serum levels for MMP3 correlate with disease activity, the best correlation with disease activity has been shown for ESR and CRP (37-39). Our

results showed significant differences in ESR and CRP between the high and low inflammatory tissue groups, but serum MMP3 levels were not significantly different. Serum MMP3 levels were correlated with the transcriptional levels of MMP3 in the synovium, thereby substantiating the idea that although increased serum MMP3 levels originates from initial production in the synovial tissues, they are not useful as soluble biomarkers reflecting tissue heterogeneity.

In the current study we did not find a clear relation between tissue inflammation status and PB gene expression markers. Previously, we demonstrated that a subgroup of RA patients has increased expression levels of type I IFN-response genes in PB cells (10). In the current study we did not find a clear relation between the expression levels of type I IFN response genes in PB cells and tissue inflammation (data not shown).

Although at the single gene level no relation between PB transcriptome profile and tissue inflammation was found, pathway-level analysis revealed that tissue subtypes correlated with differential expression of sets of genes involved in protein biosynthesis and developmental processes. Future studies are needed aimed at elucidating the possible role of these systemic biological processes in tissue inflammation. We should keep in mind that we analyzed only one affected joint which may explain the difficulties in finding specific PB cell activation markers. On the other hand, it has been shown in paired synovial biopsy samples from patients with RA that the inflammatory process in one inflamed joint is generally representative of that in other inflamed joints (20).

In summary, our results show the presence of two pathogenic processes or phases leading to disease progression: one inflammation-mediated and one immune-independent. The results highlight the importance of examination of tissue from the site of inflammation rather than merely analysis of PB.

Funding:

This study was supported by grants from The Netherlands Organization for Health Research and Development (ZonMw) in assignment of The Netherlands Organization for Scientific Research (NWO: grant number 945-02-029), the European Community's FP6 funding (Autocure), the Innovation Oriented research Program (IOP) on Genomics and the Centre for Medical Systems Biology (Netherlands Genomics Initiative). This publication reflects only the authors' views. The European Community is not liable for any use that may be made of the information herein.

Reference List

(1) Visser H, le Cessie S, Vos K, Breedveld FC, Hazes JM. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. Arthritis Rheum 2002; 46(2):357-65.

(2) Tak PP. Analyzing synovial tissue samples. What can we learn about early rheumatoid arthritis, the heterogeneity of the disease, and the effects of treatment? J Rheumatol Suppl 2005; 72:25-6.

(3) Goldblatt F, Isenberg DA. New therapies for rheumatoid arthritis. Clin Exp Immunol 2005; 140(2):195-204.

(4) Klimiuk PA, Goronzy JJ, Bjor nJ, Beckenbaugh RD, Weyand CM. Tissue cytokine patterns distinguish variants of rheumatoid synovitis. Am J Pathol 1997; 151(5):1311-9.

(5) Tak PP, Smeets TJ, Daha MR, Kluin PM, Meijers KA, Brand R et al. Analysis of the synovial cell infiltrate in early rheumatoid synovial tissue in relation to local disease activity. Arthritis Rheum 1997; 40(2):217-25.

(6) Ulfgren AK, Grondal L, Lindblad S, Khademi M, Johnell O, Klareskog L et al. Interindividual and intra-articular variation of proinflammatory cytokines in patients with rheumatoid arthritis: potential implications for treatment. Ann Rheum Dis 2000; 59(6):439-47.

(7) van der Pouw Kraan TC, van Gaalen FA, Huizinga TW, Pieterman E, Breedveld FC, Verweij CL. Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair. Genes Immun 2003; 4(3):187-96.

(8) Cantaert T, Kolln J, Timmer T, van der Pouw Kraan TC, Vandooren B, Thurlings RM et al. B lymphocyte autoimmunity in rheumatoid synovitis is independent of ectopic lymphoid neogenesis. J Immunol 2008; 181(1):785-94.

(9) Thurlings RM, Wijbrandts CA, Mebius RE, Cantaert T, Dinant HJ, van der Pouw-Kraan TC et al. Synovial lymphoid neogenesis does not define a specific clinical rheumatoid arthritis phenotype. Arthritis Rheum 2008; 58(6):1582-9.

(10) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis 2007; 66(8):1008-14.

(11) van der Pouw Kraan TC, van Gaalen FA, Kasperkovitz PV, Verbeet NL, Smeets TJ, Kraan MC et al. Rheumatoid arthritis is a heterogeneous disease: evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. Arthritis Rheum 2003; 48(8):2132-45.

(12) van der Pouw Kraan TC, van Baarsen LG, Rustenburg F, Baltus B, Fero M, Verweij CL. Gene expression profiling in rheumatology. Methods Mol Med 2007; 136:305-27.

(13) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Rustenburg F, Baggen JM, Verweij CL et al. Responsiveness to anti-TNF{alpha} therapy is related to pre-treatment tissue inflammation levels in Rheumatoid Arthritis patients. Ann Rheum Dis 2007.

(14) Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D et al. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. Nucleic Acids Res 2007; 35(Database issue): D766-D770.

(15) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006; 7(6):522-31.

(16) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-21.

(17) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-8.

(18) Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13(9):2129-41.

(19) Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res 2006; 34(Web Server issue):W645-W650.

(20) Kraan MC, Reece RJ, Smeets TJ, Veale DJ, Emery P, Tak PP. Comparison of synovial tissues from the knee joints and the small joints of rheumatoid arthritis patients: Implications for pathogenesis and evaluation of treatment. Arthritis Rheum 2002; 46(8):2034-8.

(21) Wijbrandts CA, Remans PH, Klarenbeek PL, Wouters D, van den Bergh Weerman MA, Smeets TJ et al. Analysis of apoptosis in peripheral blood and synovial tissue very early after initiation of infliximab treatment in rheumatoid arthritis patients. Arthritis Rheum 2008; 58(11):3330-9.

(22) Haringman JJ, Vinkenoog M, Gerlag DM, Smeets TJ, Zwinderman AH, Tak PP. Reliability of computerized image analysis for the evaluation of serial synovial biopsies in randomized controlled trials in rheumatoid arthritis. Arthritis Res Ther 2005; 7(4):R862-R867.

(23) Timmer TC, Baltus B, Vondenhoff M, Huizinga TW, Tak PP, Verweij CL et al. Inflammation and ectopic lymphoid structures in rheumatoid arthritis synovial tissues dissected by genomics technology: identification of the interleukin-7 signaling pathway in tissues with lymphoid neogenesis. Arthritis Rheum 2007; 56(8):2492-502.

(24) Takemura S, Braun A, Crowson C, Kurtin PJ, Cofield RH, O'Fallon WM et al. Lymphoid neogenesis in rheumatoid synovitis. J Immunol 2001; 167(2):1072-80.

(25) Young CL, Adamson TC, III, Vaughan JH, Fox RI. Immunohistologic characterization of synovial membrane lymphocytes in rheumatoid arthritis. Arthritis Rheum 1984; 27(1):32-9.

(26) Kasperkovitz PV, Timmer TC, Smeets TJ, Verbeet NL, Tak PP, van Baarsen LG et al. Fibroblast-like synoviocytes derived from patients with rheumatoid arthritis show the imprint of synovial tissue heterogeneity: evidence of a link between an increased myofibroblast-like phenotype and high-inflammation synovitis. Arthritis Rheum 2005; 52(2):430-41.

(27) van Baarsen LG, van der Pouw-Kraan TC, Verweij CL. A genomic view of subtypes in rheumatoid arthritis: towards personalized medicine. Future Rheumatology 2006; 1(3):311-22.

(28) Carlsen S, Hansson AS, Olsson H, Heinegard D, Holmdahl R. Cartilage oligomeric matrix protein (COMP)-induced arthritis in rats. Clin Exp Immunol 1998; 114(3):477-84.

(29) Morozzi G, Fabbroni M, Bellisai F, Cucini S, Simpatico A, Galeazzi M. Low serum level of COMP, a cartilage turnover marker, predicts rapid and high ACR70 response to adalimumab therapy in rheumatoid arthritis. Clin Rheumatol 2007; 26(8):1335-8.

(30) Firestein GS, Zvaifler NJ. How important are T cells in chronic rheumatoid synovitis?: II. T cell-independent mechanisms from beginning to end. Arthritis Rheum 2002; 46(2):298-308.

(31) Tak PP, Zvaifler NJ, Green DR, Firestein GS. Rheumatoid arthritis and p53: how oxidative stress might alter the course of inflammatory diseases. Immunol Today 2000; 21(2):78-82.

(32) Baeten D, Demetter P, Cuvelier C, Van den BF, Kruithof E, Van DN et al. Comparative study of the synovial histology in rheumatoid arthritis, spondyloarthropathy, and osteoarthritis: influence of disease duration and activity. Ann Rheum Dis 2000; 59(12):945-53.

(33) Nerusu KC, Warner RL, Bhagavathula N, McClintock SD, Johnson KJ, Varani J. Matrix metalloproteinase-3 (stromelysin-1) in acute inflammatory tissue injury. Exp Mol Pathol 2007; 83(2):169-76.

(34) Okada Y, Takeuchi N, Tomita K, Nakanishi I, Nagase H. Immunolocalization of matrix metalloproteinase 3 (stromelysin) in rheumatoid synovioblasts (B cells): correlation with rheumatoid arthritis. Ann Rheum Dis 1989; 48(8):645-53.

(35) van Meurs J., van LP, Holthuysen A, Lambrou D, Bayne E, Singer I et al. Active matrix metalloproteinases are present in cartilage during immune complex-mediated arthritis: a pivotal role for stromelysin-1 in cartilage destruction. J Immunol 1999; 163(10):5633-9.

(36) Murphy G, Knauper V, Atkinson S, Butler G, English W, Hutton M et al. Matrix metalloproteinases in arthritic disease. Arthritis Res 2002; 4 Suppl 3:S39-S49.

(37) Keyszer G, Lambiri I, Nagel R, Keysser C, Keysser M, Gromnica-Ihle E et al. Circulating levels of matrix metalloproteinases MMP-3 and MMP-1, tissue inhibitor of metalloproteinases 1 (TIMP-1), and MMP-1/TIMP-1 complex in rheumatic disease. Correlation with clinical activity of rheumatoid arthritis versus other surrogate markers. J Rheumatol 1999; 26(2):251-8.

(38) Roux-Lombard P, Eberhardt K, Saxne T, Dayer JM, Wollheim FA. Cytokines, metalloproteinases, their inhibitors and cartilage oligomeric matrix protein: relationship to radiological progression and inflammation in early rheumatoid arthritis. A prospective 5-year study. Rheumatology (Oxford) 2001; 40(5):544-51.

(39) Ates A, Turkcapar N, Olmez U, Tiryaki O, Duzgun N, Uguz E et al. Serum pro-matrix metalloproteinase-3 as an indicator of disease activity and severity in rheumatoid arthritis: comparison with traditional markers. Rheumatol Int 2007; 27(8):715-22.

Chapter 3.4

A subtype of multiple sclerosis defined by an activated immune defense program

3

Lisa G.M. van Baarsen[1], Tineke C.T.M. van der Pouw Kraan[1], J. Jolijn Kragt[2], Josefien M.C. Baggen[1], François Rustenburg[1], Tineke Hooper[1], Jan F. Meilof[3], Mike J. Fero[4], Christine D. Dijkstra[1], Chris H. Polman[2], Cornelis L. Verweij[1]

[1] Dept. of Molecular Cell Biology & Immunology, VUMC, Amsterdam, The Netherlands
[2] Dept. of Neurology, VUMC, Amsterdam, The Netherlands
[3] Dept. of Neurology, Martini Ziekenhuis, Groningen, The Netherlands
[4] Stanford Functional Genomics Facility, Stanford School of Medicine, Stanford, USA

Genes & Immunity 2006 Sep;7(6):522-31

Abstract

Given the heterogeneous nature of multiple sclerosis (MS), we applied DNA microarray technology to determine whether variability is reflected in peripheral blood cells. In this study we studied whole blood gene expression profiles of 29 patients with relapsing remitting MS and 25 age and sex matched healthy controls. We used microarrays with a complexity of 43k cDNAs. The data was analyzed using sophisticated pathway level analysis in order to provide insight in the deregulated peripheral immune response programs in MS.

We found a remarkable elevated expression of a spectrum of genes known to be involved in immune defense in the peripheral blood of MS patients compared to healthy individuals. Cluster analysis revealed that the increased expression of these genes was characteristic for approximately half of the patients. In addition, the gene signature in this group of patients was comparable with a virus response program.

We conclude that the transcriptional signature of the peripheral blood cells reflects the heterogeneity of MS and defines a subpopulation of relapsing remitting MS patients, who exhibit an activated immune defense program that resembles a virus response program, which is supportive for a link between viruses and MS.

Introduction

Multiple sclerosis (MS) is a chronic neurological disorder in which demyelination and inflammation occur in the white matter of the CNS (1). The disease has a heterogeneous nature, which is reflected in the clinical presentation, ranging from mild to severe demyelinating disease. The wide variation in responsiveness to treatment in MS is consistent with the heterogeneous nature of the disease. E.g. treatment with IFNbeta exhibits only partial responses in a subgroup of MS patients (2) suggesting that distinct disease mechanisms are at play in MS.

The heterogeneous nature of MS is also reflected by immunopathological studies using postmortem brain material from MS patients which revealed profound heterogeneity in the patterns of demyelination between different patients, whereas within the same patient different active plaques were similar (3). Recently, the heterogeneity of MS was also dissected *in vivo* using noninvasive MRI techniques combined with statistical remodeling (4). Although in this study the MRI data could not be correlated with pathological data, the observed subgroups were clinically meaningful. The existence of heterogeneity in the brain of MS patients is proposed to reflect different pathogenic processes underlying MS (3); an inflammation-mediated and an immune-independent demyelinating form, which both ultimately might drive brain tissue destruction. Thus, different disease mechanisms may vary among patients and, perhaps, in different stages of disease.

The heterogeneity most likely has its origin in the multifactorial nature of the disease, whereby specific combinations of environmental factors and a varying polygenic background are likely to influence not only susceptibility but also the severity and disease outcome.
Since the concordance among genetically identical twins is less than complete, an environmental factor or factors are anticipated to have a major role in the risk of developing MS in a genetically susceptible host (5). Findings from epidemiological studies, which revealed an uneven distribution of MS with disease more prevalent in temperate climates (6) as well as data from migration studies implicate an important role for an infectious agent in MS etiology (7). A number of viruses have been associated with the development of MS (8). However, despite this knowledge the environmental factor(s) that contribute to MS remain to be determined.

A powerful way to provide insight in the complexity and pathogenesis of MS has arisen from DNA microarray technology, which provides the opportunity to determine differences in gene expression of a large portion of the genome in search of genes that are differentially expressed between patients with clinically diagnosed RRMS. By large-scale gene expression profiling in blood cells from patients with MS and healthy controls one can obtain a molecular portrait that is disease associated. This molecular portrait typically represents the contributions and interactions of numerous distinct cells and diverse factors that are associated with disease. In addition, a systematic analysis of the gene expression profiles should reveal whether the heterogeneous nature of MS is reflected in the peripheral blood cells. This type of analysis will not only disclose markers that are indicative for differential disease mechanisms in MS, but may, in the case of a pathogenic stimulus of host cells, reveal a pathogen-induced transcriptional program.

Materials and Methods

Patients

Dutch patients with definite MS (9) were recruited from the outpatient clinic of the MS-Centre at the VUMC Amsterdam. We selected 29 patients with relapsing remitting (RR) disease of which 28 had not been treated with interferon beta (IFNbeta) for at least 3 years. Some of these patients were not clinically active, some had declined disease-modifying drugs (DMDs), and in others the sample was taken immediately before they started DMDs. Of the 29 patients, ten patients were included during an acute relapse and one patient who did receive IFNbeta therapy was included in the analysis. In addition, one patient was analysed twice. The control group consisted of 25 healthy volunteers, matched with respect to gender and age. All subjects were unrelated and Dutch Caucasians. This study was carried out with the approval of the Medical Ethics Committee of the VUMC, and informed consent was obtained from all subjects.

Blood sampling
2.5 ml blood was drawn in PAXgene blood RNA isolation tubes (PreAnalytix, GmbH, Germany). After blood collection the tubes were incubated for 2h at room temperature (RT) to ensure complete lyses of all the blood cells, after which the tubes were stored at -20°C. Total RNA was isolated within 7 months after blood collection. Tubes were thawed for 2 hours at RT prior to RNA isolation. Next, total RNA was isolated using the PAXgene RNA isolation kit according to the manufacturers' instructions including a DNAse (Qiagen, Venlo, The Netherlands) step to remove genomic DNA.

Sample preparation for array hybridisation
Sample preparation and array hybridisation was performed as described previously (10). Total RNA was linear amplified using the Message Amp. aRNA kit (Ambion, Huntingdon, UK), which is based on the Eberwine method (11). We used 1 μg of total RNA as input and followed the protocol according to the manufacturers' instructions allowing 5 hours for *in vitro* transcription.
Per array hybridization, two separate reaction vials were set up for labeling: one for the experimental sample to be labeled with Cy3 (green) and one for the common reference sample to be labeled with Cy5 (red). We made a common reference that consisted of a mixture of mRNAs isolated from 11 different cell lines (12) supplemented with RNA from rheumatoid synovial tissue, fibroblasts, and activated peripheral blood mononuclear cells (PBMCs). From both samples 5 μg aRNA was labeled with aminoallyl-dUTP during cDNA synthesis by reverse transcriptase, followed by chemical coupling of the aminoallyl group to Cy3 or Cy5 for the experimental and reference samples respectively.
The labeled cDNA transcripts of an MS patient or healthy control and a common reference are then hybridized together onto the 43K cDNA microarrays (Stanford University) at 65°C degrees. After 18 hours of hybridization and multiple washing steps, the hybridized cDNA products are measured by fluorescent colorimetry using the G2505B microarray scanner from Agilent Technologies (Amstelveen, The Netherlands).

Data filtering and analysis
The array images were analysed using the program GenePix Pro 3.0 (Axon Instruments Inc., Union City, CA, USA) and all data was uploaded and analysed using the Stanford Microarray Database (SMD) (http://smd.stanford.edu)(13). Intensity dependent normalisation using local estimation (14) ("Loess") was performed separately on each sector of the array ("stratified" by print-tip). In addition, "scale" normalisation was applied which makes the data more comparable across different arrays. Next we excluded data from spots with inconsistent hybridization signals (defined by a regression $R2<0.6$ for a linear fit between the Cy3 and Cy5 pixel intensities) or spots that had been flagged as defective by visual inspection during

data extraction (for instance due to scratches or dust particles). In addition, we only included spots if the signal intensity was 2.5 times higher than the background and less then the saturated value of 65000. We corrected for array batch differences by applying Single Value Decomposition (SVD). Because a single gene can be represented more than once on an array we collapsed our data by averaging the results from sequences with the same identifier (Unigene Cluster ID) in the database. Finally, we re-filtered our data in the SMD and median centered the data for genes and arrays.

We used the log2 of the background-subtracted, normalized ratio of the mean Cy5 and mean Cy3 expression values and applied several algorithms (see flowchart figure 1) for the analysis. For all algorithms we only used data from arrays if there was at least 80% good data present per gene. Significance Analysis for Microarrays (SAM) was used (15) to determine the significantly differential expressed genes between different subject groups. We considered genes as significantly differential expressed, if the False Discovery Rate (FDR) was less than 5% (a q-value of less than 0.05). Cluster (16) was used to define clusters of co-ordinately regulated genes after which Treeview visualized the clustered data.

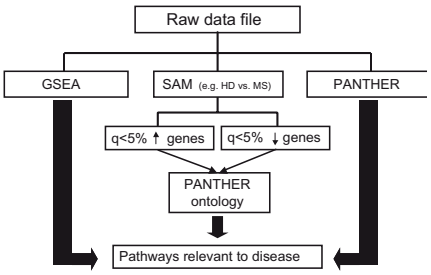


Figure 1. Flowchart data analysis
This figure shows our data analysis steps from raw microarray data towards defining the pathways relevant for disease.

Additionally, we applied PANTHER (**P**rotein **AN**alysis **TH**rough **E**volutionary **R**elationships Classification System; PE Applied Biosystems, Foster City, CA; http://PANTHER.appliedbiosystem.com (17)) to interpret our data after SAM analysis. This analysis is based on a binomial distribution test to classify the list of significantly up- or downregulated genes (referred to as "SAM gene list") in different functional categories (molecular function, biological process or biological pathway). After classification a significant p-value ($p<0.05$) indicates that a given category may be of biological interest.

Pathway-level analysis in order to find coordinated shifts in gene expression levels

We used two different algorithms both designed to find weak coordinated shifts in gene expression data: PANTHER (Applied Biosystems) and GSEA (18) (Gene Set Enrichment Analysis).

For the PANTHER pathway-level analysis we used all differentially expressed genes with their associated fold changes between two groups e.g. RRMS patients and

healthy controls. The Mann-Whitney *U*-test is used to look for coordinated changes in gene expression within each available pathway. As a result a pathway may be found significant even if the fold changes are very small.

In contrast to the PANTHER analysis, the pathway-level analysis program GSEA uses the total raw data file consisting of the normalized ratio's per gene per array. GSEA combines a priori defined sets of genes to increase signal to noise ratio and improve statistical power to determine if genes of a given gene set are enriched among the most differentially expressed genes between patients and healthy controls. We applied GSEA as described before (19), only to confirm the PANTHER analysis, therefore the data is not shown.

Results

Differential gene expression in peripheral blood (PB) of RRMS patients compared to healthy controls

We studied differences in gene expression in peripheral blood (PB) cells derived from healthy controls and patients with RRMS. A schematic outline of the data analysis is shown in Figure 1. At first the data was analyzed as two class, unpaired data using Significance Analysis of Microarrays (15) (SAM). A total of 2025 genes were selected whose transcript levels showed significantly differential expression between patients and controls. Of these genes 1059 were significantly upregulated, whereas 966 genes were downregulated in RRMS (figure 2A).

The genes with the most upregulated (>1.3 fold) expression levels encode for immunoglobulin related proteins e.g. Immunoglobulin lambda constant 2 (IGLC2) chain, Immunoglobulin lambda variable group (IGLV@) and Immunoglobulin kappa constant chain (IGKC). Other genes with marked increased expression levels in the PB cells of MS patients belonged to a group of Interferon (IFN)-induced genes e.g. induced protein 44 (IFI44L), IFN-induced transmembrane protein 1 (9-27; IFITM1), IFNalpha-inducible protein (clone IFI-15K; G1P2), IFN-induced transmembrane protein 3 (1-8U; IFITM3), Myxovirus (influenza virus) resistance 1 IFN-inducible protein p78 (mouse; Mx1), IFN-induced transmembrane protein 2 (1-8D; IFITM2), Adenosine deaminase, RNA-specific (ADAR1), IFN regulatory factor 2 (IRF2), IRF7, 2'-5'-oligoadenylate synthetase 2, 69/71kDa (OAS2) and others. Of note were also the increased levels of the antioxidant enzyme Glutathione peroxidase 1 (GPX1) and the NADPH oxidase Neutrophil cytosolic factor 1 (NCF1) that are involved in the oxidative stress response of the cell.

The top ten of the most downregulated genes encode for TAF9 RNA polymerase II (TAF9), Catalase (CAT), Myomesin (M-protein) 2 (MYOM2), CREBBP/EP300 inhibitor 1 (CRI1), Zwilch (FLJ10036), T-cell leukemia/lymphoma 1A (TCL1A), Short coiled-coil protein (SCOC), Hemoglobin, gamma G (HBG2), Claudin 12 (CLDN12), GrpE-like 2, and mitochondrial (E. coli; GRPEL2), which are involved in several different processes.

Ontology analysis of significantly differential expressed genes

To systematically categorize the 2025 significantly differential expressed genes into functional groups we applied the PANTHER ontology classification system (Applied Biosystems). PANTHER uses the binomial statistics tool (20) to compare our gene list to a reference list (NCBI: Homo sapiens genes) to determine the statistically significant over-representation of functional groups of genes. The differentially upregulated genes represented seventeen significant ontology/functional groups. These are: Alzheimer's disease, presenilin signaling, JAK/STAT signaling, Fas signaling, IFN-gamma signaling, inflammation mediated immunity, apoptosis signaling, angiogenesis, Parkinson disease, interleukin signaling, heteromeric G-protein signaling pathway-rod outer segment phototransduction, metabotropic glutamate receptor group III biology, ionotrophic glutamate receptor biology, ubiquitin proteasome signaling, D2/D3/D4 dopamine receptor mediated signaling, Toll-like receptor (TLR) signaling and cysteine biosynthesis (figure 2B). Analysis of the downregulated gene list revealed two significant categories: p53 signaling and again ubiquitin proteasome signaling (figure 2C). Since this analysis revealed that both the significantly upregulated genes as well as the downregulated genes could be grouped into ubiquitin proteasome signaling, we took a detailed look at the individual genes that constitute these categories. This revealed that most of the upregulated ubiquitin proteasome-related genes encoded proteins with ubiquitin conjugating activities to mark cellular proteins for degradation by the proteasome. In contrast, the downregulated ones represent genes that encode proteasome subunits and the UCHL5 gene encoding a deubiquitinating enzyme. Hence, the differential regulation of genes involved in ubiquitination is compatible with an increased ubiquitin proteasome protein degradation activity.

Pathway-level analysis of differentially expressed genes

Next we performed a pathway level analysis, which is based on the premise that subtle coordinate changes in gene expression levels across the whole network of genes within a certain pathway can already have major effects on the outcome. At the individual gene level, these gene expression changes do not necessary have to be significantly different. In order to find such coordinated shifts of gene expression values we applied the PANTHER (21) and GSEA (18) pathway-level algorithms that are based on the usage of the whole array data set.

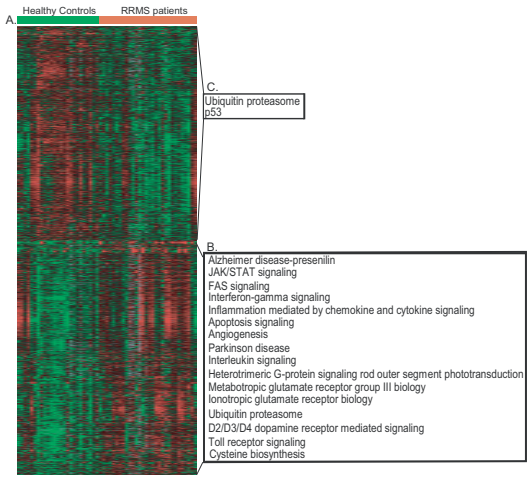
Application of PANTHER pathway-level analysis, which is restricted to pathway sets from PANTHER resources, revealed nine pathways within the RRMS patient group (table 1). These results confirmed the activation of genes involved in inflammation, IFN-gamma signaling, Alzheimer disease, JAK-STAT signaling, TLR signaling and the metabotropic glutamate receptor group III pathway, all of which were also observed in the PANTHER ontology analysis (figure 2). In addition, this analysis revealed the coordinated upregulation of genes involved in glycolysis and (nor)adrenaline biosynthesis and the downregulation of genes involved in transcriptional

regulation. These data indicate that the observed host response profile in RRMS represents the concerted action of multiple transcriptional regulators, of which a part may reflect TLR-induced innate immune activation.

Figure 2. Statistical analysis combined with PANTHER's ontology

A. Supervised cluster analysis of significantly differential expressed genes

A total of 2025 significantly differential expressed genes (FDR<5%) were selected after SAM analysis between RRMS patients and healthy controls. These genes were supervised (one-way) hierarchical clustered and visualized by Treeview. This type of clustering will place genes that show a correlated expression level in adjacent rows. Each column represents the data of one array and each row shows the relative expression levels of a single gene for all samples. Red colours mean a relative higher expression level than the median expression level across all samples. Green colours mean a relative lower expression level than the median expression level across all samples. Black colours indicate that the expression level is equal to the median expression level and grey represents missing values



B. PANTHER ontology analysis on the upregulated genes

Using the PANTHER ontology analysis the significantly upregulated genes could be classified into different functional categories.

C. PANTHER ontology analysis on the downregulated genes.

Idem as for B, using the downregulated genes.

Table 1. PANTHER pathway-level analysis: RRMS patients versus healthy controls

Pathway co-ordinately changed in RRMS:	Up/down in RRMS	In PANTHER ontology list
Inflammation mediated by chemokine and cytokine signalling pathway	Up	Yes
Alzheimer disease-presenilin pathway	Up	Yes
IFN-gamma signalling pathway	Up	Yes
Toll receptor signalling pathway	Up	Yes
JAK/STAT signalling pathway	Up	Yes
Metabotropic glutamate receptor group III pathway	Up	Yes
General transcription regulation	Down	No
Glycolysis	Up	No
Adrenaline and noradrenaline biosynthesis	Up	No

PANTHER pathway-level analysis was used to define those pathways of which the corresponding gene expression levels are co-ordinately changed throughout a significant part of the pathway. Pathways were considered significantly changed with a p-value <0.05.

We confirmed these results using the recently available Gene Set Enrichment Analysis (GSEA) pathway-analysis algorithm (18), which has the advantage that one can choose which gene sets to use for the analysis including custom-designed gene-sets (data not shown). The data from these two approaches unanimously

demonstrate the occurrence of an upregulated expression of genes involved in IFN signaling. The significantly differential expression of the IFN-response program was confirmed by real-time PCR analysis of the IFN regulated genes MxA, IFITM1 and G1P2 (data not shown). Hence, pathway-level analyses reveal that RRMS peripheral blood cells harbor features that are indicative for immune activation of TLRs possibly by microbial agents, leading to increased expression of IFNs.

Comparison MS host response program with the common pathogen response program

Table 2. Deregulated pathways in RRMS partly overlap with a common pathogen response pathway

Significant upregulated pathways in common pathogen response(22)	MS vs. HD
Inflammation mediated by chemokine and cytokine signalling pathway	X
Apoptosis signalling pathway	X
Toll receptor signalling pathway	X
Interleukin signalling pathway	X
T cell activation	
Angiogenesis	X
B cell activation	
Oxidative stress response	
EGF receptor signalling pathway	
FAS signalling pathway	X
TGF-beta signalling pathway	
PDGF signalling pathway	
Integrin signalling pathway	
Parkinson disease	X
Axon guidance mediated by semaphorins	
Heterotrimeric G-protein signalling pathway-Gq alpha and Go alpha mediated pathway	
Alzheimer disease-presenilin pathway	X
Plasminogen activating cascade	
Blood coagulation	
JAK/STAT signalling pathway	X
FGF signalling pathway	
Cadherin signalling pathway	
Purine metabolism	
Ras Pathway	
Wnt signalling pathway	
Heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha mediated pathway	
Glycolysis	X
Endothelin signalling pathway	
IFN-gamma signaling pathway	X

The common pathogen response genes described by Jenner and Young were classified into functional categories using the PANTHER's classification system. A comparison was made between the common pathogen response program and the deregulated pathways we found in RRMS. The overlapping pathways are given in bold and marked with an X. A significant part (p=0.03 Chi-square test) of the RRMS pathways overlap with the common pathogen response.

In order to create a framework for a more detailed comparison of the RRMS specific pathways with those of a microbial response program we used data from a meta-analysis of gene expression profiles after microbial infection (22). In this analysis the transcriptional profile of 32 *in vitro* studies that involved 77 different host-pathogen interactions have been analyzed together, which resulted in a defined common host-transcriptional-response program. In order to analyze which pathways are represented in the common response program we applied PANTHER pathway level analysis (table 2). This analysis revealed a total of 29 significant pathways in the common pathogen response program. Comparison of these pathways with the ones we identified in RRMS showed that there is a significant (Chi-square test p=0.03) overlap. Of the 20 RRMS specific pathways 11 are present within the common host response program, suggesting that the RRMS profile may reflect, in part, a response to a microbial infection.

Selective upregulation of the type-I IFN signaling pathway in RRMS

IFNs, which are best known for their anti-viral properties (23), are grouped into two categories, type-I and type-II IFNs. Type-I IFNs are mainly produced directly after viral infection whereas type-II IFNs are secondary produced by activated T and NK cells. Moreover, although type-I and type-II IFN response programs share a number of genes, their activities can be distinguished based on a number of differentially activated response genes. Thus a detailed analysis of the observed IFN-induced genes in RRMS is likely to disclose information on the inducing type of IFN. Therefore we obtained small but specific type-I IFN (IFNalpha; 9 genes) and type-II IFN (IFNgamma; 6 genes) gene sets from the Biocarta database (www.biocarta.com; supplementary data). To investigate the relative contribution of either gene set in the RRMS gene expression profile we calculated for each gene set the mean gene expression level (log2 ratio) per patient and healthy control and compared the two groups with each other (figure 3). This analysis showed that the mean gene expression level of the type-I IFN gene set was significantly higher in the MS patient group when compared to the healthy control group. The mean gene expression level of the type-II IFN genes was similar between patients and controls. Hence, we conclude that type-I IFNs are induced which can cause a positive feedback loop resulting in more expression of type-I IFN-induced genes.

Since our data revealed a deregulated expression of genes involved in innate immunity, like TLR and type-I IFN-response genes we studied the innate response program in more detail. Further insight in the regulatory components of the gene modules that are activated in RRMS is likely to disclose information of the stimulating component. Genes that are most strongly upregulated via TLRs are NFkB response genes via the MYD88-dependent route and IFN-stimulated genes (ISGs) via the transcription factors IRF3 and IRF7 (24). Although activation of different TLRs upregulate many of the same genes, these receptors also mediate stimulus-specific responses. Genes that are strongly upregulated via TLR4 are the ISGs and NFkB response genes. By contrast, the TLR3, TLR7/8 and TLR9 induce a pronounced

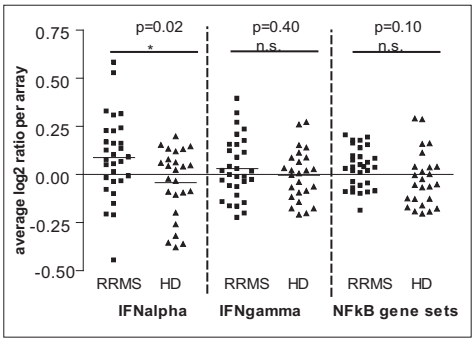


Figure 3. Increased expression of type-I IFN-induced genes in RRMS patients

The average log2 expression ratio of two IFN specific (type-I and II) gene sets as well as an NFkB specific gene set was calculated for each individual patient and control. The mean expression ratio of the type-I IFN specific genes was significantly higher in the RRMS patient group when compared to the healthy control group. In contrast, the expression ratio of type-II IFN and NFkB genes did not differ between patients and controls.

* is considered significant using an unpaired T-test; n.s. as not significant

activation of type-I ISGs, whereas the upregulation of the NFkB response gene set is relatively weak (22). Hence, the identification and relative activation of NFkB and IFN type-I specific genes in MS might provide insight in the mechanism of immune activation. In this case we used the type-I IFN specific gene set that we have used before, together with an NFkB specific gene set (23 genes) from the Biocarta database (supplementary data).

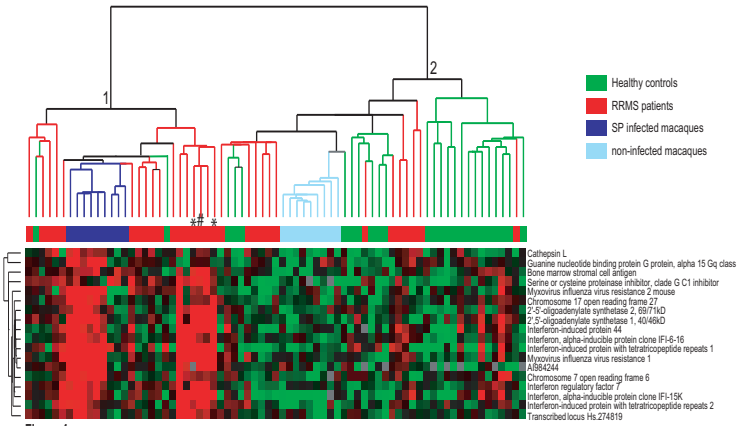
In order to confirm and extend the knowledge on the differential gene expression programs that are activated upon differential TLR stimulation, we analyzed previously reported gene expression datasets from experiments in which cells were stimulated with either a bacterial (TLR 2 and 4) or a viral (TLR 3 and 7) stimulus (22). As expected, we confirmed that bacteria especially induce the NFkB gene set whereas viral agents show a more pronounced induction of the type-I IFN specific gene set (supplementary data). To investigate the relative contribution of either gene set in the RRMS gene expression profile we calculated for both gene sets the mean gene expression level per patient and healthy control and compared the two groups with each other (Figure 3). This analysis again showed that the mean gene expression level of the type-I IFN-induced gene set was significantly higher in the MS patient group when compared to the healthy control group. However, the mean gene expression level of the NFkB specific genes was similar between patients and controls. The weak expression of NFkB genes in combination with the upregulation of type-I IFN genes would argue against a direct contribution of a TLR2 or 4 induced activation on peripheral blood cells. Therefore, these data suggest that if a microbial infection is directly involved in the peripheral blood host response in MS this is most likely of viral origin.

Gene expression profiles compared to poxvirus infection model

Next we evaluated the relationship between a viral response program and RRMS in more detail, by comparing the gene expression profiles of MS patients with the profiles of an *in vivo* virus-infection model. For that purpose we used a dataset from a nonhuman primate model for smallpox infection which provides a detailed picture of the *in vivo* host transcriptional response program in PB cells after a viral

infection (25). The data was generated on similar human cDNA microarrays that we have used, and was also stored and analyzed using the SMD. We extracted microarray data from 9 macaques before and 2 or 3 days after infection from the public access website of the SMD. These data were filtered (see M&M) and used to construct a smallpox (SP) virus-induced gene set. Therefore, we applied SAM analysis to select the significantly differential expressed genes between the PB gene expression profiles prior and after smallpox infection. The resulting 1307 statistically differential expressed genes (FDR<5%) were used for an unsupervised hierarchical clustering of the PB gene expression data from the macaques (prior and after infection), RRMS patients and healthy controls (figure 4). As anticipated, this analysis clearly separated the profiles of the macaques before and after viral infection. In addition, almost all the healthy controls clustered in the same group as the unaffected animals. Most interestingly, the MS patients were clearly divided into two groups on the basis of their gene expression profile. More than 50% of the RRMS patients clustered together with the virus-infected animals and showed an upregulation of immune defense genes. We refer to this group as the RRMS^{high} group. The patient that was analyzed twice, clustered together within this group showing the reliability of the analysis. The remaining patients, designated as the RRMS^{low} group, clustered with the uninfected animals. These results demonstrate a remarkable similarity between the virus-induced gene expression profiles of the macaques and the PB profile of a significant part (p=0.0002 Chi-Square test) of the MS patients.

Figure 4. A significant part of the RRMS patients co-clusters with smallpox (SP) infected macaques.
Genes were pre-selected of which the gene expression levels were significantly (using SAM: FDR<5%) changed in the smallpox-infected macaques when compared to the non-infected macaques. Subsequently, these genes were used for unsupervised (two-way) hierarchical cluster analysis on the combined data from the smallpox study and our study. This type of clustering will not only place genes that have correlated expression levels in adjacent rows, but also places samples that have similar expression profiles in adjacent columns. A subsequent selection of genes was made, by using only those genes whose gene expression levels were two fold different from the median expression level in at least four samples. In this figure only part of the total cluster, representing IFN-related genes, is shown. The patient that was analysed twice (marked *), clustered together in same arm of the dendogram. The patient receiving IFNbeta treatment (marked #) co-clustered with the virus-infected macaques.



Heterogeneity of RRMS

The above analysis revealed a striking heterogeneity between RRMS patients. Analysis of the overall peripheral blood profiles of only the RRMS patients using unsupervised hierarchical clustering confirmed the sub-classification of MS patients into two groups. The result of the cluster analysis correlated with the sub-classification based on the smallpox virus profile (data not shown). These results point towards an important discriminating role for immune defense activity in RRMS. To further investigate the biological basis of RRMS^{high} group we compared the expression profiles of only the RRMS^{high} group of patients with the healthy controls by PANTHER pathway-level analysis. In accordance with the overall RRMS group versus healthy controls comparison (table 1 and figure 2) using PANTHER analyses, an upregulation of genes involved in IFN signaling, Alzheimer disease, inflammation, angiogenesis, interleukin signaling, JAK/STAT signaling, TLR signaling, glycolysis and apoptosis was observed in the RRMS^{high} group. In addition, this comparison revealed the upregulation of genes involved in the Ras and EGF receptor-signaling pathway. Interestingly, all the pathways that are specific for the RRMS^{high} group overlap with those of the common response pathway from Jenner and Young (22), which is in favor of an infectious origin for the immune defense program observed in this specific patient subgroup (table 2).

Table 3. Pathway-level analyses; RRMS^{high} group versus healthy controls

Pathways significantly upregulated in RRMS^{high}:
Inflammation mediated by chemokine and cytokine signalling pathway
Angiogenesis
IFN-gamma signalling pathway
Interleukin signalling pathway
JAK/STAT signalling pathway
Toll receptor signalling pathway
Ras Pathway
Alzheimer disease-presenilin pathway
Glycolysis
EGF receptor signalling pathway

Comparison between the RRMS^{high} patient group and the healthy control group using PANTHER pathway-level analysis. Pathways were considered significant if p-value<0.05.

GSEA analysis was again used to confirm the results. Interestingly, using this analysis, especially the type-I IFN-induced and viral-response gene sets were highly significant enriched in the subgroup of RRMS (data not shown).

Clearly, the molecular heterogeneity that we observed in the PB profiles of RRMS patients is indicative for the heterogeneous nature of RRMS and may suggest the existence of different subtypes of RRMS. The question then arose as to whether the molecular stratification of RRMS was associated with clinical differences between the patients. To address this possibility, the relationship of the subtypes with

clinical and demographic parameters was determined. However, we were not able to demonstrate an association between the molecular subtypes and clinical and demographic parameters, such as age, disease duration, relapse rate and EDSS.

Discussion

Several investigators studied gene expression levels in PB cells to address the question, whether disease characteristics are detectable systemically. These previous studies demonstrated changes in PB gene expression patterns with among these genes involved in the p53 and apoptosis pathway (26-30). Whereas these studies provided insight in the genes that were differentially expressed between MS patients and healthy controls, the issue of transcript-based disease heterogeneity was not addressed.

Using pathway analysis we found next to the p53 and apoptosis pathways several other pathways and biological processes that are significantly different between RRMS patients and healthy controls. The observed upregulation of GPX1 and NCF1 together with the downregulation of catalase is suggestive for a disturbed oxidative stress response in RRMS. Although oxidative stress has been implicated to be involved in virus-induced neuronal injury (31), our results seem to be in conflict with a recent paper published by Hultqvist and co-workers (32). They showed that a mutation in the NCF1 gene, leading to a reduced oxidative burst, enhanced the autoimmune susceptibility for arthritis and encephalomyelitis in animal models. The difference may be due to a different role of oxidative burst in disease initiation versus established disease.

Another pathway that was upregulated in MS is glycolysis, which is an important metabolic pathway for the ATP biosynthesis of cells. Induction of glycolysis is typically observed in virally infected cells. Viruses do not generate the ATP that they require and rely heavily on the host cells for ATP biosynthesis.

Interestingly, we discovered for the first time significantly increased gene expression levels for several genes involved in (type-I) IFN-mediated immunity in at least half of the MS patients. One of these genes is ADAR1 (dsRNA dependent adenosine deaminase), which is induced by type-I IFNs and is involved in editing of viral RNAs (33) and the mRNA of glutamate receptor units (34). Other IFN-induced genes that showed upregulated expression levels in RRMS were OAS1, OAS2, MX1 and MX2. The OAS genes encode for dsRNA-dependent synthetases that activate the endoribonuclease RNAse L to degrade ssRNA (35). The IFN-inducible Mx proteins are GTPases that interfere with virus replication, probably by inhibiting the trafficking (36) or activity of virus polymerases (37). The increased expression of MxA confirms data from a previous report that also showed increased expression of MxA at the protein level in PB cells of RRMS patients (38). Accordingly, Chieux and colleagues (38) also showed increased levels of IFNα in the serum of MS patients.

An upregulation of IFN-induced genes is also observed in other autoimmune diseases like SLE (39), rheumatoid arthritis (40), Sjögren's syndrome (SS) (41), type I diabetes mellitus (42) and dermatomyositis (43). Moreover, the disease activity of SLE patients was correlated with increased serum levels of IFNα (44). However, for MS the available clinical data suggest that IFNbeta is involved in inhibiting the disease symptoms. These findings suggest that type-I IFNs might have differential clinical effects depending on the disease. Although in SLE the pDC in the PB have been identified as a major source of IFNα, in SS numerous IFNα -producing cells were detected in the affected salivary gland biopsies. These studies suggest that in SS the source of IFN can be from cells in the PB blood as well as from cells in the diseased tissue. Further studies are needed to demonstrate the source of IFN in MS.

Additionally, our data analysis showed a coordinated upregulation of genes involved in TLR-signalling, indicative of an activated innate immune signalling. In theory, infectious agents or endogenous factors, such as viruses, bacteria, unmethylated CpG DNA, or single- or double stranded RNA could be proximal mediators of type-I IFN production and more downstream activation programs. Innate immune responses to pathogens are believed to be patterned and stereotyped. The specialization of different TLRs allows the host to tailor its response to different pathogens. This is achieved by the differential activation of at least two different but partly overlapping pathways: the MyD88-dependent and the IRF3/7 MyD88-independent pathway. Both pathways lead to induction of the NFkB response program and gene activation via the IFN-response element (ISRE) resulting in the transcription of IFN-related genes. In general, bacterial pathogens and specific endogenous factors such as heat shock proteins induce a strong NFkB response along with a moderate or weak IFN response. In contrast, other factors such as single- or double stranded RNA or viruses induce a strong type-I IFN response along with a weak to absent NFkB response. Our results showed a high expression of type-I IFN-induced genes in MS patients, while the expression of NFkB specific genes was comparable with that of the controls. Thus if activation of PB cells of MS patients is responsible for the observed activation of IFN-induced genes, it is unlikely that this occurred via the MyD88-dependent pathway since we did not observe a difference in NFkB specific gene expression levels. Together with the observation that our data did demonstrate the upregulation of genes involved in TLR and type-I IFN-signaling, this would restrict the activation of IFN-induced genes to TLR receptors that selectively signal via the IRF3/7 pathway, like TLR3, TLR7/8, TLR9 (45) or via the recently discovered RNA helicase RIG-I that bind cytoplasmic dsRNA (46) or other thus far unknown ligands.

Infectious agents have long been considered as possible triggers for autoimmune responses. The search for a specific virus in MS may be hampered due to the fact

that there might not be a single MS-specific virus, but a variety of distinct viruses or a group of related viruses that are involved (47). Upon infection with a pathogen, cells undergo a marked reprogramming of their transcriptome (22). Part of the response program has been interpreted as a generic "alarm signal" to infection and involves functional groups of genes involved in e.g. the interferon (IFN)-response, inflammation, apoptosis and TLR signaling (22). Here we demonstrated that the pathways that were found to be different between healthy controls and RRMS patients overlap for a great part with the pathways that are induced during the common pathogen response program described by Jenner and Young (22) (table 1). This overlap is even more convincingly demonstrated in the RRMS^{high} patient group, which showed a high expression of immune defense and interferon-related genes. Moreover, we observed a striking similarity of the PB gene expression signature in MS with that of the virus infected macaque model, which demonstrates similarity of gene clusters that constitute a general "alarm signal" for a viral infection. These findings might suggest that a viral agent could play a role in MS.
However, despite the clear relation of a virus response program and the MS gene signature, for SLE and SS it was shown that immune complexes from patients IgG together with nucleic acid released by necrotic of apoptotic cells can induce type-I IFN production by PB plasmacytoid DCs and a concomitant type-I IFN response signature in the PB (48). Thus, although the data from epidemiological studies and the similarity of the PB gene expression signature in MS with that of the virus infected macaque model support a role for viruses in MS, we cannot formally exclude a contribution for endogenous factors in the induction of the type-I IFN response program in MS. Hence, further studies are needed to determine whether the increased expression of type-I IFN response genes are MS are the result of endogenous or infectious factors.

Our results clearly indicate that there is a striking difference in gene expression profiles between RRMS patients. The wide variation in clinical symptoms, disease course and therapy response is consistent with this observation. The question is what the molecular classification into RRMS^{high} and RRMS^{low} means in clinical terms. Based on the expression profiles it suggests that different pathogenic mechanisms may contribute to disease. Therefore this type of analysis provides a rational to understand the molecular and biological basis of disease heterogeneity in MS. So far we could not demonstrate an association with clinical parameters. However, we realize that the design of this study does not allow any firm conclusions to be drawn.
Because the therapeutic response to IFNbeta therapy also seems to be heterogeneous (2), it is tempting to speculate that either one of the subtypes is more likely to respond better to IFNbeta treatment. Not so surprisingly, several reports have demonstrated an induction of the type-I IFN response program in patients treated with IFNbeta (49;50). As anticipated, the one patient receiving IFNbeta therapy clustered together with the RRMS^{high} patients and showed an upregulation

of IFN-induced genes. Although transcription-based prediction of IFNbeta therapy has been studied (51;52), future research in a large cohort of patients is needed to sort out whether the absence or presence of a type-I IFN response in the peripheral blood of patients with RRMS is a prognostic factor for successful treatment with IFNbeta.

Acknowledgements

This study was partly supported by the Centre for Medical Systems Biology (a centre of excellence approved by the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research), and grants from the Stichting MS Research. We would like to thank the following authors for making their microarray data publicly available: RG Jenner (Cambridge, USA), KH Rubins (Stanford, USA) and Q Huang (Cambridge, USA).

Reference List

(1) Hafler DA. Multiple sclerosis. J Clin Invest 2004; 113(6):788-794.

(2) Rio J, Nos C, Tintore M, Borras C, Galan I, Comabella M et al. Assessment of different treatment failure criteria in a cohort of relapsing-remitting multiple sclerosis patients treated with interferon beta: implications for clinical trials. Ann Neurol 2002; 52(4):400-406.

(3) Lucchinetti C, Bruck W, Parisi J, Scheithauer B, Rodriguez M, Lassmann H. Heterogeneity of multiple sclerosis lesions: implications for the pathogenesis of demyelination. **Ann Neurol 2000; 47(6):707-717.**

(4) Bielekova B, Kadom N, Fisher E, Jeffries N, Ohayon J, Richert N et al. **MRI as a marker for disease heterogeneity in multiple** sclerosis. Neurology 2005; 65(7):1071-1076.

(5) Sadovnick AD, Armstrong H, Rice GP, Bulman D, Hashimoto L, Paty DW et al. A population-based study of multiple sclerosis in twins: update. Ann Neurol 1993; 33(3):281-285.

(6) Pugliatti M, Sotgiu S, Rosati G. The worldwide prevalence of multiple sclerosis. Clin Neurol Neurosurg 2002; 104(3):182-191.

(7) Kurtzke JF. Epidemiology of MS. In: Hallpike JF, Adams CWM, Tourtellote WE, editors. Multiple Sclerosis. Baltimore, Maryland: Willimas and Wilkins; 2005. 49-95.

(8) Ter Meulen V., Katz M. The proposed viral etiology of multiple sclerosis and related demyelinating diseases. In: Raine CS, MacFarland HF, Tourtellotte WW, editors. Multiple sclerosis. Clinical and Pathogenetic Basis ed. London: Chapman and Hall; 1997. 287-305.

(9) McDonald WI, Compston A, Edan G, Goodkin D, Hartung HP, Lublin FD et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. **Ann Neurol 2001; 50(1):121-**127.

(10) van der Pouw Kraan TC, van Baarsen LG, Rustenburg F, Baltus B, Fero M, Verweij CL. **Gene expression profiling in** rheumatology. Methods Mol Med 2007; 136:305-327.

(11) Van Gelder RN, von Zastrow ME, Yool A, Dement WC, Barchas JD, Eberwine JH. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A 1990; 87(5):1663-1667.

(12) Perou CM, Jeffrey SS, van de RM, Rees CA, Eisen MB, Ross DT et al. **Distinctive gene expression patterns in human** mammary epithelial cells and breast cancers. Proc Natl Acad Sci U S A 1999; 96(16):9212-9217.

(13) Ball CA, Awad IA, Demeter J, Gollub J, Hebert JM, Hernandez-Boussard T et al. The Stanford Microarray Database

accommodates additional microarray platforms and data formats. Nucleic Acids Res 2005; 33(Database issue):D580-D582.

(14) Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 2002; 30(4):e15.

(15) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-5121.

(16) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-14868.

(17) Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13(9):2129-2141.

(18) Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102(43):15545-15550.

(19) Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003; 34(3):267-273.

(20) Cho RJ, Campbell MJ. Transcription, genomes, function. Trends Genet 2000; 16(9):409-415.

(21) Mi H, Lazareva-Ulitsky B, Loo R, Kejariwal A, Vandergriff J, Rabkin S et al. The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Res 2005; 33(Database issue):D284-D288.

(22) Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. Nat Rev Microbiol 2005; 3(4):281-294.

(23) Samuel CE. Antiviral actions of interferons. Clin Microbiol Rev 2001; 14(4):778-809, table.

(24) Moynagh PN. TLR signalling and activation of IRFs: revisiting old friends from the NF-kappaB pathway. Trends Immunol 2005; 26(9):469-476.

(25) Rubins KH, Hensley LE, Jahrling PB, Whitney AR, Geisbert TW, Huggins JW et al. The host response to smallpox: analysis of the gene expression program in peripheral blood cells in a nonhuman primate model. Proc Natl Acad Sci U S A 2004; 101(42):15190-15195.

(26) Achiron A, Gurevich M, Friedman N, Kaminski N, Mandel M. Blood transcriptional signatures of multiple sclerosis: unique gene expression of disease activity. Ann Neurol 2004; 55(3):410-417.

(27) Aune TM, Maas K, Moore JH, Olsen NJ. Gene expression profiles in human autoimmune disease. Curr Pharm Des 2003; 9(23):1905-1917.

(28) Maas K, Chan S, Parker J, Slater A, Moore J, Olsen N et al. Cutting edge: molecular portrait of human autoimmune disease. J Immunol 2002; 169(1):5-9.

(29) Mandel M, Gurevich M, Pauzner R, Kaminski N, Achiron A. Autoimmunity gene expression portrait: specific signature that intersects or differentiates between multiple sclerosis and systemic lupus erythematosus. Clin Exp Immunol 2004; 138(1):164-170.

(30) Satoh J, Nakanishi M, Koike F, Miyake S, Yamamoto T, Kawai M et al. Microarray analysis identifies an aberrant expression of apoptosis and DNA damage-regulatory genes in multiple sclerosis. Neurobiol Dis 2005; 18(3):537-550.

(31) Valyi-Nagy T, Dermody TS. Role of oxidative damage in the pathogenesis of viral infections of the nervous system. Histol Histopathol 2005; 20(3):957-967.

(32) Hultqvist M, Olofsson P, Holmberg J, Backstrom BT, Tordsson J, Holmdahl R. Enhanced autoimmunity, arthritis, and encephalomyelitis in mice with a reduced oxidative burst due to a mutation in the Ncf1 gene. Proc Natl Acad Sci U S A 2004; 101(34):12646-12651.

(33) Patterson JB, Thomis DC, Hans SL, Samuel CE. Mechanism of interferon action: double-stranded RNA-specific adenosine deaminase from human cells is inducible by alpha and gamma interferons. **Virology 1995; 210(2):508-511.**

(34) Lomeli H, Mosbacher J, Melcher T, Hoger T, Geiger JR, Kuner T et al. **Control of kinetic properties of AMPA receptor channels** by nuclear RNA editing. Science 1994; 266(5191):1709-1713.

(35) Player MR, Torrence PF. The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation. Pharmacol Ther 1998; 78(2):55-113.

(36) Reichelt M, Stertz S, Krijnse-Locker J, Haller O, Kochs G. Missorting of LaCrosse virus nucleocapsid protein by the interferon-induced MxA GTPase involves smooth ER membranes. Traffic 2004; 5(10):772-784.

(37) Stranden AM, Staeheli P, Pavlovic J. Function of the mouse Mx1 protein is inhibited by overexpression of the PB2 protein of influenza virus. Virology 1993; 197(2):642-651.

(38) Chieux V, Chehadeh W, Hautecoeur P, Harvey J, Wattre P, Hober D. Increased levels of antiviral MxA protein in peripheral blood of patients with a chronic disease of unknown etiology. J Med Virol 2001; 65(2):301-308.

(39) Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. **Proc Natl Acad Sci U S A 2003; 100(5):2610-2615.**

(40) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. **Rheumatoid arthritis** subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. **Ann Rheum Dis 2007; 66(8):1008-1014.**

(41) Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML et al. **Activation of the type I interferon system in** primary Sjogren's syndrome: a possible etiopathogenic mechanism. Arthritis Rheum 2005; 52(4):1185-1195.

(42) Huang X, Yuang J, Goddard A, Foulis A, James RF, Lernmark A et al. Interferon expression in the pancreases of patients with type I diabetes. Diabetes 1995; 44(6):658-664.

(43) Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R et al. Interferon-alpha/beta-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005; 57(5):664-678.

(44) Bengtsson AA, Sturfelt G, Truedsson L, Blomberg J, Alm G, Vallin H et al. Activation of type I interferon system in systemic lupus erythematosus correlates with disease activity but not with antiretroviral antibodies. Lupus 2000; 9(9):664-671.

(45) Akira S, Takeda K. Toll-like receptor signalling. Nat Rev Immunol 2004; 4(7):499-511.

(46) Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M et al. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. **Nat Immunol 2004; 5(7):730-737.**

(47) van Noort JM, Bajramovic JJ, Plomp AC, van Stipdonk MJ. **Mistaken self, a novel model that links microbial infections with** myelin-directed autoimmunity in multiple sclerosis. J Neuroimmunol 2000; 105(1):46-57.

(48) Lovgren T, Eloranta ML, Bave U, Alm GV, Ronnblom L. Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG. **Arthritis** Rheum 2004; 50(6):1861-1872.

(49) Weinstock-Guttman B, Badgett D, Patrick K, Hartrich L, Santos R, Hall D et al. **Genomic effects of IFN-beta in multiple** sclerosis patients. J Immunol 2003; 171(5):2694-2702.

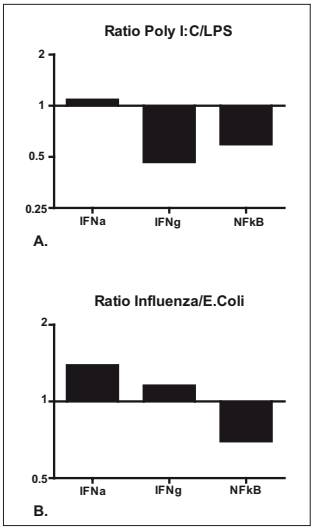
(50) Koike F, Satoh J, Miyake S, Yamamoto T, Kawai M, Kikuchi S et al. Microarray analysis identifies interferon beta-regulated genes in multiple sclerosis. J Neuroimmunol 2003; 139(1-2):109-118.

(51) Baranzini SE, Mousavi P, Rio J, Caillier SJ, Stillman A, Villoslada P et al. Transcription-based prediction of response to IFNbeta using supervised computational methods. PLoS Biol 2005; 3(1):e2.

(52) Sturzebecher S, Wandinger KP, Rosenwald A, Sathyamoorthy M, Tzou A, Mattar P et al. Expression profiling identifies responder and non-responder phenotypes to interferon-beta in multiple sclerosis. Brain 2003; 126(Pt 6):1419-1429.

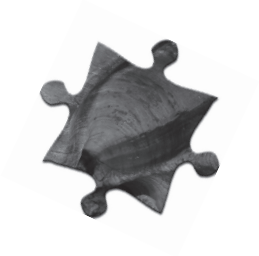
Supplementary Table S1.

Biocarta specific gene sets:		
IFNalpha	IFNgamma	NFkB
IFNA1	IFNG	CHUK
IFNAR1	IFNGR1	FADD
IFNAR2	IFNGR2	IKBKB
IFNB1	JAK1	IKBKG
ISGF3G	JAK2	IL1A
JAK1	STAT1	IL1R1
STAT1		IRAK1
STAT2		MAP3K1
TYK2		MAP3K14
		MAP3K7
		MAP3K7IP1
		MYD88
		NFKB1
		NFKBIA
		RELA
		RIPK1
		TLR4
		TNF
		TNFAIP3
		TNFRSF1A
		TNFRSF1B
		TRADD
		TRAF6



Supplementary Figure 1.

Gene expression data from the review of Jenner & Young was used in order to determine the relative contribution of type-I, type-II and NFkB specific gene sets (from Biocarta database) in bacterial- (represented by LPS and E. Coli) and viral- (represented by Poly I:C and Influenza) induced gene expression profiles. For each gene set the average log2 ratio was calculated for dendritic cells that were in vitro stimulated with either PolyI:C, LPS, Influenza or E. Coli. Subsequently the ratio's between PolyI:C and LPS (A) and between Influenza and E.Coli (B) was plotted into a graph. Figure A shows that LPS stimulation induces more expression of IFNgamma and NFkB genes than Poly I:C stimulation. In contrast, Poly I:C stimulation induces more expression of IFNalpha specific genes. Figure B shows that E.Coli stimulation induces more expression of NFkB specific genes than Influenza, whereas Influenza stimulation induced more expression of IFNalpha and gamma specific genes.



Chapter 4. Pharmacogenomics

Chapter 4.1
Pharmacogenomics of infliximab treatment using peripheral blood cells of patients with rheumatoid arthritis

4

Lisa G.M. van Baarsen[1,2*], Carla A. Wijbrandts[2*], Danielle M. Gerlag[2], François Rustenburg[1], Tineke C.T.M. van der Pouw Kraan[1], Ben A.C. Dijkmans[1,3], Paul P. Tak[2], Cornelis L. Verweij[1]

[1]VU University Medical Center, Amsterdam, Netherlands; [2]Academic Medical Center/ University of Amsterdam, Amsterdam, Netherlands; [3]Jan van Breemen Institute, Amsterdam, Netherlands

** Both authors contributed equally*

Submitted for publication

Abstract

Objective: The goal of this study was to provide insight into the pharmacological changes in the peripheral blood molecular profile induced by TNF blockade in patients with rheumatoid arthritis (RA).

Methods: Peripheral blood was obtained in PAXgene tubes from 33 RA patients before and one month after TNF-blocking therapy (infliximab). From pre- and post-treatment samples gene expression profiles were determined from 15 randomly chosen patients out of the 33 patients and the remaining 18 RA patients served as validation cohort. Hierarchical clustering, Significance Analysis of Microarrays and PANTHER pathway-level analysis were used for data-analysis and interpretation. From genes of interest the expression levels were confirmed by quantitative real-time PCR.

Results: A group-based paired analysis of the gene expression profiles from the post- vs. pre-treatment samples revealed a signature of genes significantly regulated by TNF blockade. The most significantly down-regulated gene was Peptidase Inhibitor 3 (PI3/SKALP), which expression level was down-regulated 1.84-fold after treatment. The down-regulated genes reflected several biological pathways such as inflammation, angiogenesis, B and T-cell activation. Further analysis revealed that the pharmacological response signature was significantly regulated in all treated patients, irrespective of clinical response which is indicative for the presence of an active TNF pathway in all RA patients.

Conclusions: The results favor a model for the parallel presence of TNF-dependent and TNF-independent pathways in the individual RA patient. Clinical response status to TNF blockade may be dependent on the relative contribution of TNF-independent effector pathways.

Introduction

Tumor necrosis factor alpha (TNF) is a proinflammatory cytokine that plays a critical role in the pathogenesis of chronic inflammatory autoimmune diseases like rheumatoid arthritis (RA).(1) Accordingly, TNF antagonists have proven to be effective for the treatment of RA.(2;3) Blockade of TNF reduces the acute phase response and decreases both local and systemic levels of inflammatory mediators in patients with RA (reviewed in (4)). However, the improvement varies between patients, and approximately 30% of RA patients fail to respond to this therapy. Clinical studies revealed that the response to treatment ranges from an almost complete remission to even worsening of clinical symptoms.

The variation in the primary clinical response to treatment suggests that different effector pathways contribute to disease activity between patients. This could be either the consequence of the existence of parallel pathways in a single patient whereby the relative contribution of TNF-independent effector pathways varies

between patients, irrespective of the level of available TNF-bioactivity. Alternatively, the variation in response to treatment could be explained by differences in bioactive TNF levels associating a good response to the presence and a failure to respond to the absence of bioactive TNF. This model is supported by the observation that indirectly measured levels of TNF-bioactivity were predictive of clinical response to TNF-blockade.(5) In addition, pre-treatment expression levels of TNF in synovial tissue is higher in responders compared to non-responders.(6) This would implicate that the failure to respond is associated with the presence of a TNF-independent pathway with concomitant lack of a TNF-pathway.

In normal physiology TNF achieves its biological effects by binding to multi-subunit receptor TNFR1 (p50) or TNFR2 (p75) on the cell surface. Soluble TNF (sTNF) binds with higher affinity to TNFR1 (7) while transmembrane TNF (mTNF) primarily signals through TNFR2.(8) Whereas TNFR1-mediated signaling is well known for its pro-inflammatory and apoptosis-inducing capacities the signaling through TNFR2 is less well understood. Moreover, mTNF by itself may lead to intracellular signaling (9), making the TNF-induced signaling program even more complex. Ultimately the signaling cascade leads to transcriptional activation of TNF-responsive genes. Likewise, the capacity of TNF-specific blockers to modulate gene expression can be considered as the qualitative and quantitative measure of the inhibition of TNF bioactivity. Thus, investigating changes in gene transcripts after adding or deleting a certain mediator of the immune response reflects the transcriptional response to that mediator.

With the aid of genomics technology, we are now in a position to determine the genome wide pharmacological outcome of TNF blockade in RA patients. We hypothesize that an in-depth understanding of the transcriptional changes after TNF-blockade in RA will allow us to evaluate the presence and/or absence of TNF-induced pathways at the individual patient level. Therefore, we measured the pharmacological effect of TNF-blockade in RA patients using large-scale gene expression profiling of peripheral blood (PB) cells.

▪ Methods

Patients

Thirty-three consecutive patients with RA according to the ACR criteria were enrolled in the study from the outpatient clinic of the Academic Medical Center (AMC) in Amsterdam over a period of one year. Inclusion criteria were: 18-85 years of age, a failure to respond to at least two disease modifying anti-rheumatic drugs (DMARDs) including methotrexate (MTX), and active disease (defined as having a disease activity score evaluated in 28 joints (DAS28) ≥ 3.2). Patients with a history of an acute inflammatory joint disease of different origin or previous use of a TNF

blocking agent were excluded. Patients were on stable maximally tolerable MTX treatment. Whole blood samples (2.5 ml) were obtained using PAXgene tubes (PreAnalytix, GmbH, Germany) from 33 RA patients prior to initiation of TNF blocking therapy with infliximab (3 mg/kg intravenously, giving according to the usual dosing schedule for RA at baseline, at week 2, 6, and subsequently every 8 weeks). A second PAXgene tube was obtained after one month of treatment. All patients gave written informed consent and the study protocol was approved by the Medical Ethics Committee (AMC). Gene expression profiles were determined from 15 randomly chosen patients out of the 33 patients. The remaining 18 patients were used as a validation group to confirm gene expression levels by quantitative real-time PCR. An overview of the patients' characteristics is given in Table 1. After 16 weeks of treatment the clinical response to treatment was assessed using both the EULAR criteria (10;11) as well as the reduction in DAS28 of at least 1.2.(12)

Table 1. Baseline patients' characteristics†

	Array samples (n=15)	qPCR samples (n=18)
Age (years)	51 (39-55)	58 (51-69)
Gender (female/male)	7/8	14/4
Disease Characteristics:		
DAS28	5.6 (4.6-7.0)	5.7 (5.0-6.6)
CRP (mg/dl)	8 (6-22)	13 (5-44)
ESR (mm/hour)	25 (12-41)	32 (16-47)
ACPA titer (U/ml)	100 (15-595)	541 (121-1805)
IgM RF titer (U/ml)	28 (14-133)	67 (22-182)
Disease duration (months)	77 (29-240)	65 (36-1992)
Erosions	n=13	n=15
Medication:		
MTX dose (mg/week)	25 (20-30)	21 (15-25)
Prednisone use	n=2	n=5
NSAID use	n=7	n=12

† Values are listed as median (interquartile range 25-75) unless indicated otherwise; DAS28, 28-joint Disease Activity Score; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ACPA, anti-citrullinated peptide antibodies; RF, Rheumatoid Factor; MTX, methotrexate; NSAID, Non-Steroidal Anti-Inflammatory Drugs

Blood sampling for RNA isolation

Blood (2.5 ml) was drawn in PAXgene blood RNA isolation tubes (PreAnalytix) and stored at -20°C. RNA was isolated using the PAXgene RNA isolation kit according to the manufacturer's instructions including a DNAse (Qiagen, Venlo, Netherlands) step to remove genomic DNA. Quantity and purity of the RNA was tested using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, DE).

Sample hybridization for microarray analysis

We used 43K cDNA microarrays from the Stanford Functional Genomics Facility (http://microarray.org/sfgf/) printed on aminosilane-coated slides containing

~20.000 unique genes. Only one batch of arrays was used for all experiments. First DNA spots were UV-cross linked to the slide using 150-300 mJoules. Sample preparation and microarray hybridization were performed as described previously.(13;14)

Microarray data analysis

Data storage and filtering was performed using the Stanford Microarray Database (15) (http://genome-www5.stanford.edu//) as described previously.(16) Raw (log2) data can be downloaded from the publicly accessible Stanford database website. To determine the significant pharmacological response to treatment we performed a paired analysis using significance analysis of microarrays (SAM).(17) A gene was considered as significantly differential expressed if the False Discovery Rate (FDR) was equal to or less than 5%. Cluster analysis (18) was used to define clusters of co-coordinately changed genes after which the data was visualized using Treeview. To interpret our data after SAM analysis, we applied PANTHER (Protein ANalysis THrough Evolutionary Relationships) Classification System (Applied Biosystems, Foster City, CA) at http://PANTHER.appliedbiosystem.com. (19;20) This analysis uses the binomial statistics tool to compare the list of significantly up- or down-regulated genes to a reference list (NCBI Homo sapiens) to statistically determine over- or underrepresentation of PANTHER classification categories such as biological processes. A Bonferroni correction was applied to correct for multiple testing and after classification analysis a significant P-value ($P < 0.05$) indicates that a given category may be of biological interest.

Real-time PCR

In 48-wells PCR plates (Greiner Bio-One GmbH, Frickenhausen, Germany) 0.5 μg of total RNA was reverse transcribed into cDNA using the Revertaid H-minus cDNA synthesis kit (MBI Fermentas, St. Leon-Rot, Germany) according to the manufacturer's instructions. Quantitative real-time PCR was performed using an ABI Prism 7900HT Sequence detection system (Applied Biosystems) using SybrGreen (Applied Biosystems). Primers were designed using Primer Express software and guidelines (Applied Biosystems). To calculate arbitrary values of mRNA levels and to correct for differences in primer efficiencies a standard curve was constructed. Expression levels of target genes were expressed relative to 18SRNA.

Statistical analysis

Pre- and post-treatment values were compared by paired t-test analysis using Graphpad Prism 4 (Graphpad Software, Inc., La Jolla, CA) Differences were considered statistically significant with P-values less than 0.05.

Results

Pharmacogenomics response to infliximab treatment

The bioactivity of TNF in RA patients can be measured retrospectively by analyzing the change in gene transcription after TNF blockade. In this explorative study we set out to determine the transcriptional response to TNF-blocking therapy in RA using DNA microarray analysis of peripheral blood (PB) cells. Therefore, gene expression profiles of 15 RA patients before (t=0) and one month after (t=1) TNF-blocking therapy (infliximab) were compared. Using a false discovery rate (FDR) of <5% SAM revealed 1623 significantly regulated genes: 692 genes were significantly up-regulated whereas 931 genes were significantly down-regulated after one month of treatment (for total gene list see supplementary Table S1). Accordingly, these genes encompass the pharmacogenomic anti-TNF response signature and are referred to as such in this article.

The most significantly down-regulated gene was Peptidase Inhibitor 3 (PI3/SKALP), which expression level was down-regulated 1.84-fold after treatment. This significant down-regulation was confirmed by quantitative real-time PCR in the independent validation cohort of 18 RA patients receiving infliximab treatment ($p=0.02$; data not shown). To verify that the gene expression changes were indeed regulated by TNF-blockade, two different gene sets involved in TNFR1 and TNFR2 signaling were obtained (http://www.biocarta.com). For all genes in both gene sets the average value in expression level for each patient was calculated and pre-treatment levels were compared with post-treatment levels. This revealed a significant down-regulation of both gene sets after TNF-blockade (paired t-test $P < 0.05$, data not shown) confirming that both TNFR signaling pathways are affected by infliximab.

To unravel the biological function of the pharmacogenomic anti-TNF response signature we applied PANTHER's classification analysis. The down-regulated genes could be classified into several biological pathways such as inflammation, angiogenesis, and B and T-cell activation (Table 2) indicating that blocking TNF dampens these immune-related pathways. The up-regulated genes consisted of many genes with unknown function and could therefore not be classified into pathways.

Collectively, these data demonstrate the existence of a universal pharmacogenomic anti-TNF response signature reflecting the pharmacological changes in PB after neutralization of TNF.

Table 2. PANTHER pathway-level analysis of significantly differential expressed genes

Pathways down-regulated genes:	Ref. list	FDR<5%	Expected	P-value
Angiogenesis	229	21	6.11	2.32E-04
Inflammation mediated by chemokine and cytokine signaling pathway	315	25	8.41	3.18E-04
B cell activation	86	12	2.3	7.18E-04
Integrin signaling pathway	227	20	6.06	7.30E-04
T cell activation	111	13	2.96	1.86E-03
FGF signaling pathway	148	15	3.95	2.23E-03
PDGF signaling pathway	189	17	5.05	2.82E-03
TGF-beta signaling pathway	154	15	4.11	3.50E-03
Ras Pathway	91	11	2.43	6.39E-03
EGF receptor signaling pathway	150	14	4	1.01E-02
Endothelin signaling pathway	98	11	2.62	1.22E-02
Alzheimer disease-amyloid secretase pathway	77	9	2.06	3.93E-02

FDR, False Discovery Rate; Ref. list, NCBI Homo sapiens reference gene list

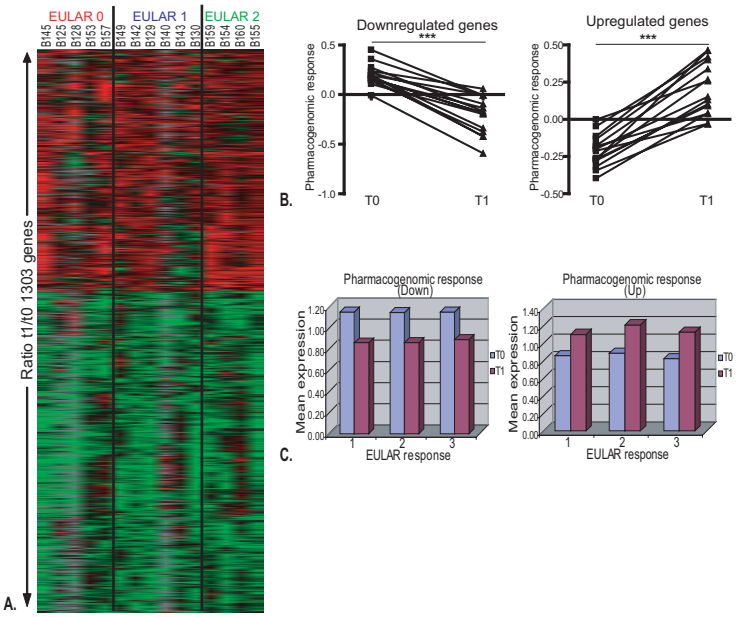
Pharmacogenomic anti-TNF response signature is common in all patients

The identification of a pharmacogenomic anti-TNF response signature allows evaluation of the presence or absence of a TNF-blockade response for each individual patient. This knowledge will provide information regarding the presence or absence of TNF bioactivity in relation to clinical response. Therefore, we calculated for each patient the treatment-induced change (ratio T=1/T=0) in transcript level for each gene in the pharmacogenomic anti-TNF response signature and visualized the response values between patients using supervised one-way hierarchical cluster analysis (Figure 1A). This analysis revealed significant changes in the expression level of the pharmacogenomic anti-TNF response signature in all RA patients tested. Thus all patients carry features of TNF bioactivity irrespective of clinical response.

However, this analysis does not give much information about the possible inter-individual difference in the extent of pharmacological response. Hence, for each individual patient a mean expression value for both the up- and down-regulated pharmacogenomic response genes was calculated and pre-treatment values were compared with post-treatment values (Figure 1B). These results revealed that all patients treated displayed a similar magnitude of the pharmacogenomic anti-TNF response signature.

Next we wanted to analyze if inter-individual differences present at baseline or after treatment may be related to clinical response to treatment. Therefore, for each clinical response group (poor, moderate and good responders using EULAR criteria) the average expression levels for both the up- and down-regulated pharmacogenomic anti-TNF response genes were calculated and the different response groups were compared with each other (Figure 1C). Baseline expression

Figure 1. Pharmacogenomic response to anti-TNF treatment



One-way (supervised) hierarchical cluster analysis using the 1623 pharmacogenomic response genes whose expression levels were significantly (FDR<5%) changed after one month of TNF blockade. Genes were only included if both baseline and post-treatment data were present in at least 80% of the patients, resulting in a total selection of 1303 genes. Up-regulated genes after therapy are indicated by a red color, down-regulated by a green color and genes that show no differences in expression after therapy are indicated in black. Figure 1B shows the individual pre- vs. post-treatment comparison of the mean expression values of the pharmacogenomic anti-TNF response signature. Figure 1C shows that the extent of pharmacogenomic response as well as the baseline and post-treatment levels are not related to EULAR response (1=poor; 2=moderate; 3=good). *** $P < 0.001$ (paired T test)

levels, pharmacological change as well as remaining expression levels for the genes encompassing the above described pharmacogenomic anti-TNF response signature were similar between the different clinical response groups. Collectively, these data could not reveal consistent differences in the pharmacogenomic anti-TNF response signature between clinical responders and non-responders.
These results imply that all RA patients have a certain level of the active TNF response program prior to treatment, which is reflected by a common pharmacological response towards TNF-blockade irrespective to clinical response to treatment.

Heterogeneity in pharmacological anti-TNF response

Whereas all patients showed a common pharmacogenomic response, not all patients exhibited a good clinical response to treatment. This suggests that other pathways might contribute to disease, which may include pathways that are differentially affected by TNF-blockade therapy. To explore the existence of subtle inter-individual differences in the pharmacodynamic anti-TNF response, we searched for differences between the pharmacological anti-TNF responses at the individual patient level by calculating for each patient and for each gene the ratio of gene expression pre- vs. post-therapy (log-2 ratios). This allowed us to search for gene expression differences that are missed in the group (post- vs. pre-

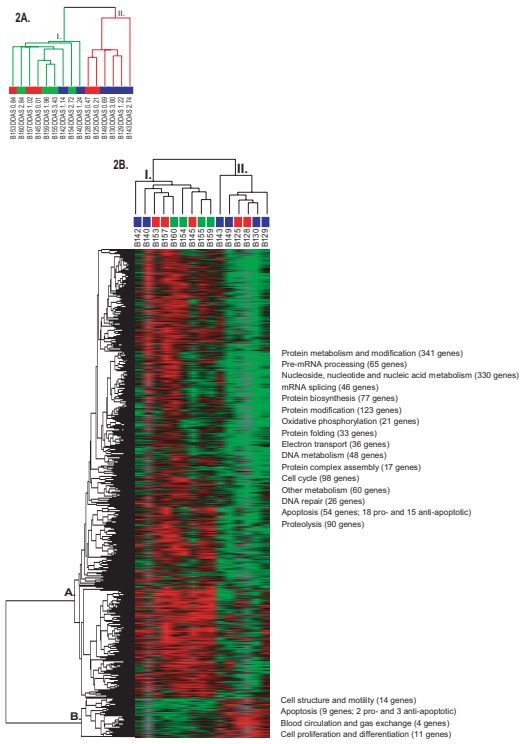


Figure 2. Inter-individual differences in treatment-induced changes in gene expression

Gene expression ratio's (biological response) that differed at least two-fold after treatment in at least 4 patients were two-way hierarchically clustered. The resulting 440 genes divided the RA patients into two different patient groups (I and II) based on their difference in pharmacological response (A). SAM revealed 2777 significantly (FDR<5%) differential changed genes between groups I and II. Cluster analysis was used to visualize these significant pharmacological differences in gene expression changes and PANTHER classification analysis was applied to determine the biological relevance of the pharmacological differences (B). Up-regulated genes after therapy are indicated by a red color, down-regulated by a green color and genes that show no differences in expression after therapy are indicated in black. At the top the EULAR response is given in colors: Red, 0= poor response; blue, 1=moderate response; green, 2=good response.

treatment)-based comparison. A total of 440 genes that showed at least a two-fold change after treatment in at least 4 patients were found. Unsupervised two-way hierarchical clustering based on the expression levels of these 440 genes divided the RA patients into two different patient groups (I and II) (Figure 2A).

Usage of different criteria for the gene selection generally resulted in an identical subdivision, indicative for the robustness of this cluster-based subclassification of patients (data not shown). SAM on all the measured genes revealed 249 genes that were increased and 2528 genes that were decreased in group II (FDR<5%; Figure 2B). PANTHER classification analysis revealed that the up-regulated 2528 genes in group I are involved in processes like oxidative phosphorylation, DNA repair, (pro-) apoptosis, protein metabolism, modification and biosynthesis, whereas 249 genes up-regulated in group II are involved in (anti-)apoptosis, cell proliferation and cell differentiation.

These analyses performed at the individual patient level demonstrate heterogeneity in the biological processes that are activated as a result of treatment. The observed variation in the pharmacological response might help to understand (part) of the mechanism explaining the differential clinical responsiveness to TNF blockade.

Discussion

The current study demonstrates the pharmacological effect of TNF blockade on the PB transcriptome profile. Four weeks after the start of TNF blockade significant changes in the expression levels were found. Pathway level analysis of these significant changes in gene expression levels revealed an overall dampening of immune response after TNF blockade. According to the pro-inflammatory cytokine activity of TNF we observed a down-regulation of genes involved in inflammation, angiogenesis, T and B cell activation after TNF blocking therapy. Interestingly, we observed this down-modulation of inflammation in every treated RA patient irrespective of clinical response, implying that all RA patients had an active TNF response program prior to treatment contributing to disease.

Good pharmacodynamic (PD) markers are needed to improve the prediction of drug efficacy and safety at the individual patient level. These quantitative PD markers should reflect features of drug exposure and drug response with respect to modulation of the molecular target, the cognate biochemical pathways and/or downstream biological effects.(21;22) The availability of these quantitative PD markers may provide a rational basis for decision making during e.g. treatment optimization. A relatively easy to measure PD marker currently described for TNF-blocking therapy is serum CRP levels.(2) In addition, cytokine levels for IL-1RA and IL-6 are described to be downregulated early after TNF blockade and the change in IL-6 levels is associated with CRP reduction.(23) However, most of the described PD markers are assessed by using mean levels of patient groups and therefore do not necessarily apply for each individual patient. Our findings show a significant downregulation of several pro-inflammatory genes after one month of infliximab treatment in all treated patients, irrespective of clinical response, indicating that these genes are treatment-specific. Therefore, this anti-TNF response signature provides a rich framework for selecting PD markers potentially useful for monitoring the pharmacological response of TNF-blockade therapy. Especially gene targets, such as the most significantly downregulated gene SKALP, may be selected as possible PD marker to monitor anti-TNF response. Studying the kinetics of these genes should reveal if they can also be used as early PD markers and genes encoding for a serum protein may be useful as serum protein PD markers.

Despite the highly beneficial effects of TNF blockade in suppressing disease, clinical improvement appears to be limited to a subset of patients. Clinical studies revealed that the response to treatment ranges from an almost complete remission to even worsening of clinical symptoms. The few pharmacogenomics studies published so far were conducted in order to find genes predictive for clinical response to TNF-blocking therapy.(24-28) Overall, the identification of predictive biomarkers for clinical response to treatment has not yet led to consistent results, although synovial TNF expression may explain about 10% of the effect.(6) As indicated by the

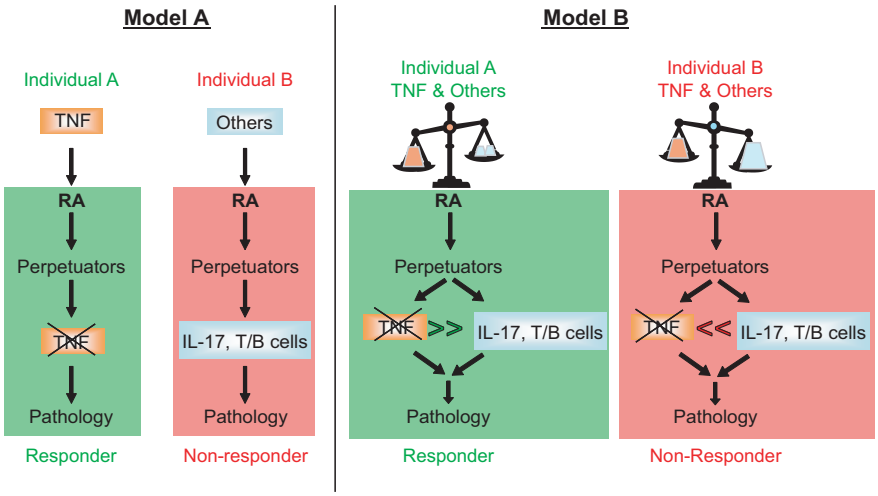


Figure 3. Models describing the possible role of TNF in disease pathogenesis of RA
The pharmacodynamic response to TNF blockade is restricted to patients who show a good clinical response to treatment (model A). In the individual RA patient the balance between TNF-dependent and independent pathways determines the treatment effect of TNF blockade (model B).

small sample size per clinical response group, the current study was not designed to identify predictors of clinical response to treatment. In contrast to the predictive studies, the present study was performed to determine the pharmacological response in PB induced by TNF-blocking therapy. Our results provide evidence for the concept that all treated patients reveal a comparable qualitative and quantitative pharmacological response upon TNF-blockade, indicative for the neutralization of bioactive TNF irrespective of clinical response. These findings would reject the model wherein the pharmacodynamic response is restricted to patients who show a good clinical response (Figure 3, model A). Instead, our results favor a model of parallel TNF-dependent and TNF-independent pathways in the individual patient, wherein the relative contribution of the TNF-independent pathways varies between patients and may determine the clinical outcome (Figure 3, model B). Candidates for TNF-independent pathways involve cells and mediators such as IL-17, B-cells and T-cells that have been shown to play major roles in the pathogenesis of RA.(29) Clinical response to TNF blockade may be determined by the contribution of the different TNF-independent pathways to disease. If the TNF pathway is dominant over the others, patients are more likely to elicit clinical improvement after TNF blockade.(6;30) Conversely, when besides TNF other pathways are either playing a major role in disease pathogenesis or are activated as a result of treatment, patients will show no improvement of disease activity.

The overall presence of an activated TNF-pathway in all treated RA patients is in line with a study showing joint protection and changes in inflammatory markers by infliximab plus MTX treatment even in patients without clinical improvement,(31) although the interpretation of this study has been questioned by others.(32;33) In

addition, our results are complementary to previous work showing that a reduction of TNF mRNA levels in whole blood did not predict clinical response in RA patients treated with infliximab.(34)
Despite the presence of an active TNF response program in all RA patients treated, not all patients responded well to infliximab treatment suggesting that other gene transcripts may change differentially between patients. Indeed, a detailed analysis revealed the existence of subtle interindividual differences in the PD response that separated RA patients in two groups. Moreover, starting from a hypothesis-driven approach we identified differential effects in the activation of IFN-response genes upon blockade of TNF (van Baarsen et al. unpublished observations). Further studies with larger sample sizes are needed to investigate the relation between the identified different PD response patterns and the primary clinical response to treatment.

Funding

This study was supported by a grant from The Netherlands Organization for Health Research and Development (ZonMw) in assignment of The Netherlands Organization for Scientific Research (NWO) (grant number 945-02-029), the European Community's FP6 funding (Autocure), the Innovation Oriented research Program (IOP) on Genomics and the Centre for Medical Systems Biology (Netherlands Genomics Initiative). This publication reflects only the authors' views. The European Community is not liable for any use that may be made of the information herein.

Reference List

(1) Feldmann M, Brennan FM, Maini RN. Rheumatoid arthritis. Cell 1996; 85(3):307-310.

(2) Elliott MJ, Maini RN, Feldmann M, Kalden JR, Antoni C, Smolen JS et al. Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor alpha (cA2) versus placebo in rheumatoid arthritis. Lancet 1994; 344(8930):1105-1110.

(3) Feldmann M, Maini RN. Anti-TNF alpha therapy of rheumatoid arthritis: what have we learned? Annu Rev Immunol 2001; 19:163-196.

(4) Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. Pharmacol Ther 2008; 117(2):244-279.

(5) Marotte H, Maslinski W, Miossec P. Circulating tumour necrosis factor-alpha bioactivity in rheumatoid arthritis patients treated with infliximab: link to clinical response. Arthritis Res Ther 2005; 7(1):R149-R155.

(6) Wijbrandts CA, Dijkgraaf MG, Kraan MC, Vinkenoog M, Smeets TJ, Dinant H et al. The clinical response to infliximab in rheumatoid arthritis is in part dependent on pretreatment tumour necrosis factor alpha expression in the synovium. Ann Rheum Dis 2008; 67(8):1139-1144.

(7) Grell M, Wajant H, Zimmermann G, Scheurich P. The type 1 receptor (CD120a) is the high-affinity receptor for soluble tumor necrosis factor. Proc Natl Acad Sci U S A 1998; 95(2):570-575.

(8) Grell M, Douni E, Wajant H, Lohden M, Clauss M, Maxeiner B et al. The transmembrane form of tumor necrosis factor is

the prime activating ligand of the 80 kDa tumor necrosis factor receptor. Cell 1995; 83(5):793-802.

(9) Pocsik E, Duda E, Wallach D. Phosphorylation of the 26 kDa TNF precursor in monocytic cells and in transfected HeLa cells. J Inflamm 1995; 45(3):152-160.

(10) van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum 1996; 39(1):34-40.

(11) Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol 2005; 23(5 Suppl 39):S93-S99.

(12) van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998; 41(10):1845-1850.

(13) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis 2007; 66(8):1008-1014.

(14) van der Pouw Kraan TC, van Baarsen LG, Rustenburg F, Baltus B, Fero M, Verweij CL. Gene expression profiling in rheumatology. Methods Mol Med 2007; 136:305-327.

(15) Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D et al. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. Nucleic Acids Res 2007; 35(Database issue): D766-D770.

(16) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006; 7(6):522-531.

(17) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001; 98(9):5116-5121.

(18) Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998; 95(25):14863-14868.

(19) Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R et al. PANTHER: a library of protein families and subfamilies indexed by function. Genome Res 2003; 13(9):2129-2141.

(20) Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A et al. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. Nucleic Acids Res 2006; 34(Web Server issue):W645-W650.

(21) Danhof M, de Lange EC, la Pasqua OE, Ploeger BA, Voskuyl RA. Mechanism-based pharmacokinetic-pharmacodynamic (PK-PD) modeling in translational drug research. Trends Pharmacol Sci 2008; 29(4):186-191.

(22) Workman P. How much gets there and what does it do?: The need for better pharmacokinetic and pharmacodynamic endpoints in contemporary drug discovery and development. Curr Pharm Des 2003; 9(11):891-902.

(23) Charles P, Elliott MJ, Davis D, Potter A, Kalden JR, Antoni C et al. Regulation of cytokines, cytokine inhibitors, and acute-phase proteins following anti-TNF-alpha therapy in rheumatoid arthritis. J Immunol 1999; 163(3):1521-1528.

(24) Koczan D, Drynda S, Hecker M, Drynda A, Guthke R, Kekow J et al. Molecular discrimination of responders and nonresponders to anti-TNFalpha therapy in rheumatoid arthritis by etanercept. Arthritis Res Ther 2008; 10(3):R50.

(25) Lequerre T, Gauthier-Jauneau AC, Bansard C, Derambure C, Hiron M, Vittecoq O et al. Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. Arthritis Res Ther 2006; 8(4):R105.

(26) Lindberg J, af KE, Catrina AI, Nilsson P, Klareskog L, Ulfgren AK et al. Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients. Arthritis Res Ther 2006; 8(6):R179.

(27) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Rustenburg F, Baggen JM, Verweij CL et al. Responsiveness to anti-TNF{alpha} therapy is related to pre-treatment tissue inflammation levels in Rheumatoid Arthritis patients. Ann Rheum Dis 2007.

(28) de Vries N, Tak PP. The response to anti-TNF-alpha treatment: gene regulation at the bedside. Rheumatology (Oxford)

2005; 44(6):705-707.

(29) McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. Nat Rev Immunol 2007; 7(6):429-442.

(30) Miossec P. Cytokines in rheumatoid arthritis: is it all TNF-alpha? Cell Mol Biol (Noisy -le-grand) 2001; 47(4):675-678.

(31) Smolen JS, Han C, Bala M, Maini RN, Kalden JR, van der HD et al. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. Arthritis Rheum 2005; 52(4):1020-1030.

(32) Pincus T, Yazici Y, Yazici H, Kavanaugh AF, Kremer JM, Wolfe F. Radiographic benefit without clinical improvement in infliximab-treated patients with rheumatoid arthritis: comment on the article by Smolen et al. Arthritis Rheum 2005; 52(12):4044-4045.

(33) van den Berg WB, van Riel PL. Uncoupling of inflammation and destruction in rheumatoid arthritis: myth or reality? Arthritis Rheum 2005; 52(4):995-999.

(34) Pachot A, Arnaud B, Marrote H, Cazalis MA, Diasparra J, Gouraud A et al. Increased tumor necrosis factor-alpha mRNA expression in whole blood from patients with rheumatoid arthritis: reduction after infliximab treatment does not predict response. J Rheumatol 2007; 34(11):2158-2161.

Chapter 4.2

Regulation of IFN response gene activity during infliximab treatment in rheumatoid arthritis is associated with clinical response to treatment

4

Lisa G.M. van Baarsen[1], Carla A. Wijbrandts[2], François Rustenburg[1], Tineke Cantaert[2], Tineke C.T.M. van der Pouw Kraan[3], Dominique L. Baeten[2], Ben A.C. Dijkmans[4], Paul P. Tak[2], Cornelis L. Verweij[1,4]

[1]Department of Pathology, VU University Medical Center, Amsterdam, Netherlands; [2]Department of Clinical Immunology & Rheumatology, Academic Medical Center, University of Amsterdam, Netherlands; [3]Department of Molecular Cell Biology & Immunology, VU University Medical Center, Amsterdam, Netherlands; [4]Department of Rheumatology, VU University Medical Center, Amsterdam, Netherlands

Submitted for publication

Abstract

Introduction: Cross-regulation between TNF and type I IFN has been postulated to play an important role in autoimmune diseases. Therefore we determined the effect of TNF-blockade in rheumatoid arthritis (RA) on the type I IFN-response gene activity in relation to clinical response.
Methods: Peripheral blood from 33 RA patients was collected in PAXgene tubes before and after the start of infliximab treatment. In a first group of 15 patients the baseline expression of type I IFN-regulated genes was determined using cDNA-microarrays and compared to levels one month after treatment. The remaining 18 patients were used as an independent group for validation using quantitative (q)PCR.
Results: Gene expression analysis revealed that anti-TNF antibody treatment induced a significant increase in type I IFN-response activity in a subset of RA patients, whereas expression levels remained similar or were slightly decreased in others. The findings appear clinically relevant since patients with an anti-TNF induced increased IFN-response gene activity had a poor clinical outcome. This association was confirmed and extended for an IFN-response gene set consisting of *OAS1*, *LGALS3BP*, *Mx2*, *OAS2* and *SERPING1* in five EULAR good and five EULAR poor responders, by qPCR.
Conclusions: Regulation of IFN-response gene activity upon TNF-blockade in RA is not as consistent as previously described, but varies between patients. The differential changes in IFN-response gene activity appear relevant to the clinical outcome of TNF-blockade in RA.

4

Introduction

Cytokines are key regulators of pathogenic processes in a variety of inflammatory and autoimmune diseases. Major roles for both TNF and type I interferon (IFN) have previously been demonstrated. Type I IFN (IFNα/β) play an important role in systemic lupus erythematosus (SLE) (1). Evidence for the role of IFN in SLE came from the induction of disease during IFNα/β treatment and circulating IFN-inducers (2;3). Type I IFN activity in SLE is associated with disease severity (1). TNF was the first cytokine convincingly demonstrated to contribute to chronic inflammation in several autoimmune diseases, including rheumatoid arthritis (RA) and Crohn's disease (4). Accordingly, blockade of TNF activity has proven to be highly beneficial in the treatment of these diseases (5;6).

Blockade of TNF reduces the acute phase reaction and decreases the local and systemic levels of inflammatory mediators in patients with RA (reviewed in (7)). However, the improvement varies between patients, and approximately 30% of RA patients fail to respond to this therapy. It has been suggested that TNF suppresses IFNα production by inhibiting both the generation of plasmacytoid dendritic cells

(pDC) and their IFNα secretion (8;9). Accordingly, it was shown that TNF blockade in systemic onset juvenile idiopathic arthritis (SoJIA) patients is associated with a higher expression of IFN-response genes (9). The authors concluded that inhibition of TNF activity in SoJIA patients is associated with increased transcription of IFNα regulated genes. The expression of type I IFN response genes in the peripheral blood cells reflects a direct means to measure the *in vivo* IFN-response inducing bioactivity. Similar findings were made for patients with primary Sjögren's syndrome (SS) who were treated with a TNF antagonist (10). Here, the type I IFN bioactivity in the blood was measured in an indirect manner, based on the use of a bioassay wherein a serum sample is tested to induce the expression of IFN response activity. Together, these findings might explain why anti-TNF therapy is not effective in SLE (11) and is associated with increased anti-dsDNA autoantibody titers. Conversely, evidence is available from studies that IFNβ decreased TNF production by peripheral blood mononuclear cells in healthy individuals (12). These observations suggest reciprocal regulation between TNF and type I IFNs, which represent two opposing vectors in immune homeostasis, whose disbalance may lead to immunopathology.

Therefore, we aimed to determine the effect of TNF blockade on the type I IFN-response gene activity in RA patients in relation to the clinical response. Previously, we demonstrated increased expression levels of type I IFN response genes in approximately 50% of anti-TNF treatment naive RA patients compared to healthy controls (13). The IFN signature was unrelated to medication and disease activity. This analysis was based on the measurement the expression of 34 type I IFN response genes directly in the peripheral blood of RA patients. Accordingly, other researchers reported shared IFN response features in peripheral blood cells of a subset of RA and SLE patients (14). In concordance with these results, another report demonstrated increased levels of IFNα in serum of a subset of RA patients (15). We were interested to know whether TNF blockade in RA led to a consistent increase in type I IFN-response gene activity as was reported for SoJIA and SS. Moreover, we were interested to know whether anti-TNF-induced changes in IFN-response activity were relevant for the clinical outcome of TNF-blockade.

▪ Materials & Methods

Patients

Consecutive patients with RA according to the ACR criteria were enrolled in the study at the outpatient clinic of the Academic Medical Center (AMC) in Amsterdam over a period of one year. Inclusion criteria were: 18-85 years of age, a failure of at least two disease modifying anti-rheumatic drugs including methotrexate (MTX) and active disease (DAS28 ≥ 3.2). Patients with a history of an acute inflammatory joint disease of different origin or previous use of a TNF blocking agent were excluded. Patients were on stable, maximally tolerable MTX treatment. Whole

blood samples (2.5 ml) were obtained using PAXgene tubes (PreAnalytix, GmbH, Germany) from 33 RA patients prior to initiation of anti-TNF therapy with infliximab (3 mg/kg intravenously at baseline, week 2, 6, and subsequently every 8 weeks). After one, two, three and four months of treatment another PAXgene tube was obtained. All patients gave written informed consent and the study protocol was approved by the Medical Ethics Committee (AMC). After 16 weeks of treatment clinical response was assessed using the EULAR response criteria (16;17) as well as the reduction in DAS28 (response defined by a decrease in DAS28 ≥1.2) (18). An overview of the patients' characteristics is given in Table 1.

Table 1. Baseline patients' characteristics†

	Array analysis (n=15)	qPCR analysis (n=18)
Age	51 (39-55)	58 (51-69)
Gender (female/male)	7/8	14/4
Disease Characteristics:		
DAS28	5.6 (4.6-7.0)	5.7 (5.0-6.6)
CRP (mg/dl)	8 (6-22)	13 (5-44)
ESR	25 (12-41)	32 (16-47)
ACPA titer (U/ml)	100 (15-595)	541 (121-1805)
IgM RF titer (U/ml)	28 (14-133)	67 (22-182)
Disease duration (months)	77 (29-240)	65 (36-1992)
Erosions	n=13	n=15
Medication:		
MTX dose (mg/week)	25 (20-30)	21 (15-25)
Prednisone	n=2	n=5
NSAID	n=7	n=12

† Values are listed as median (interquartile range 25-75) unless indicated otherwise; DAS28, 28-joint Disease Activity Score; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ACPA, anti-citrullinated protein antibodies; RF, Rheumatoid Factor; MTX, methotrexate; NSAID, Non-Steroidal Anti-Inflammatory Drugs

Blood sampling for RNA isolation

2.5 ml blood was drawn in PAXgene blood RNA isolation tubes (PreAnalytix, GmbH, Germany) and stored at -20°C. Tubes were thawed for 2 hours at room temperature prior to RNA isolation. Next, total RNA was isolated using the PAXgene RNA isolation kit according to the manufacturer's instructions including a DNAse (Qiagen, Venlo, Netherlands) step to remove genomic DNA. Quantity and purity of the RNA was tested using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, DE).

Microarray data

We used 43K cDNA microarrays from the Stanford Functional Genomics Facility (http://microarray.org/sfgf/) printed on aminosilane-coated slides containing ~20.000 unique genes. Only one batch of arrays was used for all experiments. First DNA spots were UV-cross linked to the slide using 150-300 mJoules. Sample

preparation and microarray hybridization were performed as described previously (13;19). Data storage and filtering was performed using the Stanford Microarray Database (20) (http://genome-www5.stanford.edu//) as described previously (21). Raw data (log2) can be downloaded from the publicly accessible Stanford database website.

IFN-response gene set

Previously, we showed that a prominent cluster of highly correlated type I IFN-response genes is upregulated in a subgroup of biological-naive RA patients compared to healthy controls (13). A gene set consisting of 34 type I IFN-response genes was obtained from these data. A smaller IFN gene set consisting of 15 genes was selected for validation analysis using Fluidigm's BioMark™ Real-Time PCR System. Detailed information of the gene lists is given in Table 2.

Quantitative real-time PCR

RNA (0.5 μg) was reverse transcribed into cDNA using the Revertaid H-minus cDNA synthesis kit (MBI Fermentas, St. Leon-Rot, Germany) according to the manufacturer's instructions. Quantitative real-time PCR was performed using an ABI Prism 7900HT Sequence detection system (Applied Biosystems, Foster City, CA) using SybrGreen (Applied Biosystems). Primers were designed using Primer Express software and guidelines (Applied Biosystems) and used primer sequences are listed in supplementary Table S1. To calculate arbitrary values of mRNA levels and to correct for differences in primer efficiencies a standard curve was constructed. Expression levels of target genes were expressed relative to *18SRNA*.

Fluidigm's BioMark™ Real-Time PCR System

The BioMark™ 48.48 Dynamic Array (Fluidigm Corporation, San Francisco, CA) for real-time quantitative PCR was used to simultaneously measure the expression of 15 IFN-response genes (See Table 2) in 47 samples (plus 1 negative control) in triplicate. This experiment was performed at the outsourcing company ServiceXS (Leiden, Netherlands). Used pre-designed Taqman Gene Expression Assays are listed in supplementary Table S1. Expression levels of target genes were expressed relative to *18SRNA*.

Statistical analysis

Data were analyzed using software programs Graphpad Prism 4 (Graphpad Software, Inc., La Jolla, CA) and SPSS version 14.0 (Chicago, IL). First data were checked for normal (Gaussian) distribution. Paired t-test analysis was used to compare pre- and post-treatment expression levels. Two-group comparisons were analyzed using unpaired t test or two-way ANOVA analysis, where appropriate. Data were considered significant with *P*-values less than 0.05.

Table 2. IFN-response gene sets used in study

IFN set (34 genes)	Validation (15 genes)	Symbol	Name
X		AA075725	none
X		AA142842	none
X		AI347124	none
X		ATF3	Activating transcription factor 3
X		EIF2AK2	Eukaryotic translation initiation factor 2-alpha kinase 2
X	X	EPSTI1	Epithelial stromal interaction 1 (breast)
X		Hs.128576	CDNA FLJ90394 fis, clone NT2RP2005632
X		Hs.97872	Transcribed locus
X		IFI16	Interferon, gamma-inducible protein 16
X	X	IFI35	Interferon-induced protein 35
X	X	IFI44L	Interferon-induced protein 44-like
X	X	IFI6	Interferon, alpha-inducible protein 6
X	X	IFIT1	Interferon-induced protein with tetratricopeptide repeats 1
X		IFIT2	Interferon-induced protein with tetratricopeptide repeats 2
X	X	IFITM1	Interferon induced transmembrane protein 1 (9-27)
X		IL1RN	Interleukin 1 receptor antagonist
X	X	IRF2	Interferon regulatory factor 2
X		IRF7	Interferon regulatory factor 7
X	X	ISG15	ISG15 ubiquitin-like modifier
X	X	LGALS3BP	Lectin, galactoside-binding, soluble, 3 binding protein
X		MX1	Myxovirus (influenza virus) resistance 1
X	X	MX2	Myxovirus (influenza virus) resistance 2 (mouse)
X	X	OAS1	2',5'-oligoadenylate synthetase 1, 40/46kDa
X	X	OAS2	2'-5'-oligoadenylate synthetase 2, 69/71kDa
X		PARP14	Poly (ADP-ribose) polymerase family, member 14
X		PLSCR1	Phospholipid scramblase 1
IFN set (34 genes)	Validation (15 genes)	Symbol	Name
X		RNF213	Ring finger protein 213
X	X	RSAD2	Radical S-adenosyl methionine domain containing 2 (alias cig5)
X		RTP4	Receptor (chemosensory) transporter protein 4
X	X	SAMD9L	Sterile alpha motif domain containing 9-like
X	X	SERPING1	Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1
X		TAP1	Transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)
X		TNFAIP6	Tumor necrosis factor, alpha-induced protein 6
X		UBE2L6	Ubiquitin-conjugating enzyme E2L 6

Results

Differential effect of TNF blockade on type I IFN signature

Previously, we compared the gene expression profiles of peripheral blood cells of RA patients to those of healthy controls and found that a subgroup of RA patients has an increased expression of type I IFN-response genes (13). This increased expression in IFN-response genes was highly variable between the individual RA patients and unrelated medication and disease activity. In the present study we determined the gene expression profiles of peripheral blood cells of 15 RA patients before and one month after anti-TNF (infliximab) treatment. The effect of TNF blockade on the transcription of type I IFN-response genes was investigated using the type I IFN-response gene set described in our previous RA study (13). For each patient the expression levels of 34 type I IFN-response genes (Table 2) were averaged and baseline values were compared to post-treatment levels (Figure 1A). At the patient group level there was no significant change in type I IFN-response gene activity upon initiation of TNF blockade. The regulation of IFN-response genes upon TNF blockade was highly variable between patients.

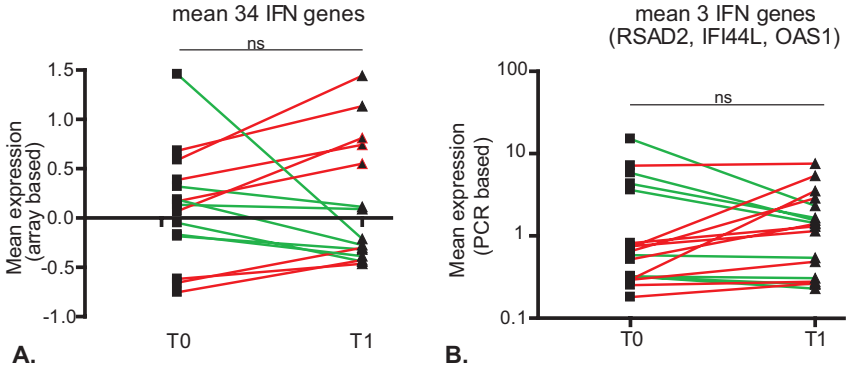


Figure 1. Differential regulation of IFN-response genes upon TNF blockade

The expression levels of 34 type I IFN-response genes were determined by cDNA microarray analysis in peripheral blood cells of 15 patients before (T0) and one month after (T1) anti-TNF treatment. Subsequently, for each patient the expression levels were averaged (note: data are in log2 space) and baseline levels were compared to post-treatment levels (A). The patients whose IFN-response gene levels are induced after TNF blockade are indicated by red lines and patients with a downregulation by green lines. Subsequently, the expression levels of three (RSAD2, IFI44L and OAS1) IFN-response genes were measured by quantitative real-time PCR in an independent group of 18 patients. The expression levels of the three genes were averaged and baseline levels were compared to post-treatment levels (B). ns: not significant using a paired t test analysis

To confirm these results by quantitative real-time (q)PCR in an independent cohort of 18 RA patients, 3 genes (*RSAD2*, *IFI44L* and *OAS1*) were selected that showed the best correlation (R>0.9) with the mean expression value of the type I IFN-response gene set used in the previous analysis. The mean expression of the 3 genes was measured by qPCR for all 18 RA patients before and one month after infliximab

therapy. Ten patients showed on an increased expression of these three IFN-response genes after TNF blockade, whereas in eight patients similar or decreased levels were observed (Figure 1B). Collectively, these results confirm findings from the microarray study and evidently demonstrate that the regulation of IFN-response gene activity upon TNF-blockade in RA is not as consistent as previously described for SLE and SS.

Change in type I IFN-response gene activity is unrelated to baseline levels

Since the type I IFN-response gene expression levels are already highly heterogeneous in biological naive RA patients, we questioned whether the observed changes were related to the magnitude of IFN-response gene expression prior to treatment. Therefore the relationship between the extent of the baseline IFN-response gene expression levels and its change after TNF blockade was tested. In the 15 patients the baseline mean expression of the type I IFN gene set did not correlate with their corresponding change after treatment (Pearson R= -0.42, p=0.12). This was confirmed in the validation group of 18 patients using the mean expression levels of the three IFN-response genes (*RSAD2*, *IFI44L* and *OAS1*) measured by qPCR, although a trend towards significance was observed (Spearman R=-0.44, p=0.064).

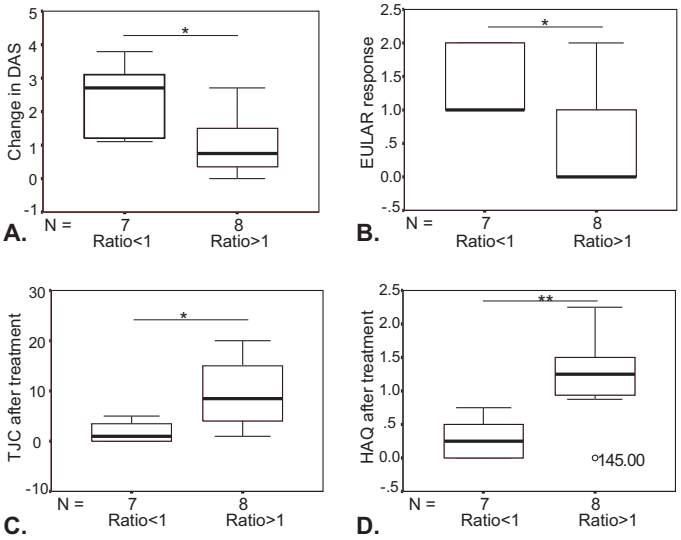
These findings reveal that the type I IFN-response gene expression profile prior to treatment is not associated with the direction of its change upon TNF blockade.

TNF/IFN cross regulation and clinical response to treatment

Finally, we investigated whether the treatment-induced changes in type I IFN-response gene expression levels were associated with clinical response to treatment. Therefore, the patients (n=15) were divided into two groups based on their change (ratio>1 and ratio<1) in mean expression level for the 34 IFN-response genes as demonstrated in Figure 1A. Subsequently clinical parameters were compared between these two groups. Clinical response to treatment was determined after 16 weeks of treatment. Interestingly, the patients that showed an increase in type I IFN-response gene expression levels after one month of treatment had a worse clinical response to treatment. This was reflected by less improvement in disease activity scores (DAS, P=0.013) and higher tender joint counts (TJC, P=0.015) and higher health assessment questionnaire-disability scores (HAQ-DI, P=0.008) after treatment (Figure 2). Accordingly, all patients without an anti-TNF induced increase in type I IFN gene activity had a good or moderate response to treatment as assessed by the EULAR response criteria (P=0.018, Figure 2).

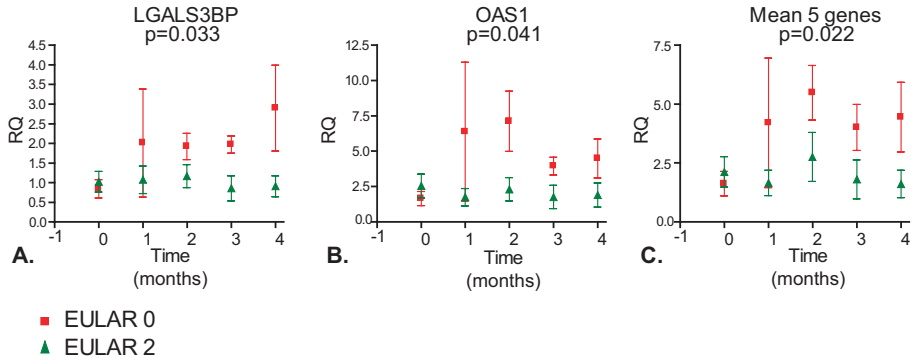
To determine whether the IFN response to TNF blockade was sustained over time, five EULAR good and five EULAR poor responders were selected and the expression levels of 15 IFN-response genes (selected from the 34 gene set used above, see Table 2) were measured at baseline, 1, 2, 3 and 4 months after treatment by qPCR.

Figure 2. Differential regulation of IFN-response genes upon TNF blockade and clinical response to treatment



Patients were divided into two groups (ratio<1 and ratio>1) based on their IFN-response upon TNF blockade and compared to each other with respect to clinical response to treatment. Ratio is determined by the T=1/T=0 expression levels of the IFN-response genes as demonstrated in Figure 1A. Data are shown as box plots; each box represents the 25th to 75th percentiles. The lines inside represents the median and the end of the whiskers represent the smallest and largest observation. Patients with an upregulation in IFN-response genes displayed a significant (unpaired t test, * P<0.05; **P<0.01) worse clinical response to treatment as assessed by change in DAS (DAS before treatment minus DAS 16 weeks after treatment, A), EULAR response (B), tender joint count (TJC) (C) and HAQ-DI (D) scores after treatment.

Figure 3. Poor response to TNF blockade is accompanied with upregulation of IFN-response genes



For five EULAR (0) poor responder (red color) and five EULAR (2) good responder (green color) patients the expression levels of 15 IFN-response genes were measured by quantitative real-time PCR (Biomark) at baseline, one, two, three and four months after treatment. From two poor and one good responder patients the three months time points are missing. The IFN-response gene expression levels during treatment were compared between the two clinical response groups using two-way ANOVA test. Treatment-induced change in the expression levels of two genes; LGALS3BP (A) and OAS1 (B) were significantly different between the two response groups. The mean expression level of five IFN-response genes (LGALS3BP, OAS1, Mx2, SERPING1 and OAS2) showed the best significant difference between the two clinical response groups. RQ means relative quantity. Graphs show the mean and SEM expression levels for each clinical response group.

The expression levels were averaged for the individual patients and the treatment-induced changes in IFN-response gene expression levels over time were compared between the two clinical response groups using two-way ANOVA. Overall, the IFN-response genes showed an upregulation in the poor responder group, which was most prominent two months after the start of therapy (data not shown). At single gene level the increased expression in poor versus good responders reached significance for the *OAS1* and *LGALS3BP* genes (Figure 3A and 3B). For three other IFN-response genes (*Mx2*, *OAS2* and *SERPING1*) a trend (~p=0.06, data not shown) was observed towards increased expression in the poor responder patients. Combining these five genes (*OAS1*, *LGALS3BP*, *Mx2*, *OAS2* and *SERPING1*) together into one IFN-response gene set improved the significance (Figure 3C). These data demonstrate that poor response to infliximab treatment is associated with treatment-induced increase in type I IFN-response gene activity.

▪ Discussion

In this explorative study we demonstrated that blockade of the inflammatory cytokine TNF in RA patients modulates the expression of IFN-response gene activity in a heterogeneous manner. A combined expression value of several type I IFN-response genes was determined to reduce the extensive overlap in individual gene expression levels that exists between responder and non-responders. The data reveal that some RA patients display a treatment-induced increased expression of type I IFN-response genes, whereas others display no effect or a small decrease in type I IFN-response gene activity. The divergent effect of TNF inhibition on expression of IFN-response genes is consistent with the heterogeneous nature of RA and the heterogeneous clinical response to TNF blocking therapy. We provided evidence that the treatment-induced change in expression levels of IFN-response genes is associated with the EULAR response rate at 16 weeks after the start of infliximab treatment. RA patients who revealed an increased IFN-response gene expression profile after one to two months of anti-TNF treatment exhibited a poor clinical response. No association between clinical response to infliximab treatment and baseline IFN-response gene activity was found.

IFNs are known for their immune regulatory properties. Previously, we provided evidence for a type I IFN-response signature in RA (13). Upregulation of type I IFN-response genes has been reported in peripheral blood cells of (a subset of) patients with other autoimmune diseases, like SLE (1;22-24), dermatomyositis (25) and multiple sclerosis (21). Type I IFNs (IFNαβ) exert broad dual effects on the immune system, reflecting both immune stimulatory and immune suppressive activities. Immune stimulatory activities relate to activation of myeloid DC, chemokines, chemokine receptors, costimulatory molecules (CD40, CD80 and CD86) and humoral responses. Immune suppressive effects are reflected by Th2 cell skewing and anti-

proliferative and pro-apoptotic effects. According to their dual effect on immunity, their role in disease may range from detrimental to beneficial. Although the anti-TNF induced increase in IFN-response activity might be an epiphenomenon related to the effect of TNF blockade, it is tempting to speculate on a role of increased IFN bioactivity in the deteriorating clinical effects. The association between an increase in type I IFN-response gene activity and poor response to anti-TNF treatment may suggest a harmful role for type I IFN bioactivity in RA or, alternatively, a failed attempt to counter regulate inflammation.

The differences between the effects of TNF blockade on IFN-response activity between the studies in SS (10) and SoJIA (9) on the one hand, and our studies in RA and spondyloarthritis (SpA) (26) on the other could have their origin in differences in design between the studies, such as the use of infliximab in RA versus etanercept in SS, and the different read out system used in the SS study which was based on an indirect reporter-cell assay to measure type I IFN activity in plasma. Using the reporter assay we previously demonstrated that the serum type I IFN bioactivity was increased in SpA patients treated with etanercept whereas it was transiently declined by infliximab (26). This result is indicative for differential effects of etanercept and infliximab on IFN-response activity. In an attempt to find an explanation for the apparent discrepancies, we learned that in both RA and SS an increased IFN-response activity is associated with a poor clinical response.
The inter-individual differences in anti-TNF induced IFN-response may be the result of differential regulatory processes. Evidence is available that TNF blockade may exert both inhibitory and activating effects on IFN response activity. *In vitro* experiments suggested that endogenous secretion of TNF by pDCs represents a negative feedback on IFN production (9). Whereas this finding suggested that TNF displays counteracting effects on IFN-response activity, others have reported that TNF initiates an IRF1-dependent autocrine loop leading to sustained expression of STAT1-dependent type I IFN-response genes (27). Hence the divergent outcome of the IFN-response activity might be a consequence of differences in the relative contribution of each of these processes in the regulation of IFN-response activity. Alternatively, genetic variation in the type I IFN biology could underlie the variation in response activity. Single nucleotide polymorphisms (SNPs) in several transcription factors involved the type I IFN pathway (e.g. *IRF5*, *Tyk2*, *STAT4*) have recently been associated with a number of autoimmune diseases including SLE (28;29) and RA (30-32). Future studies are needed to unravel the mechanism behind the divergent alterations in IFN-response gene activity upon TNF blockade.

▪ Conclusions

In summary, this study shows that there is a large variation between RA patients in the change of IFN-response gene expression levels after TNF blockade. The change

in IFN-response genes is unrelated to baseline expression levels. Interestingly, treatment-induced increase of IFN-response gene activity is associated with poor clinical response to infliximab treatment. Additional studies in larger patient cohorts should reproduce and confirm these findings.

Funding

This study was supported by grants from The Netherlands Organization for Health Research and Development (ZonMw) in assignment of The Netherlands Organization for Scientific Research (NWO: grant number 945-02-029), the European Community's FP6 funding (Autocure), the Innovation Oriented research Program (IOP) on Genomics and the Centre for Medical Systems Biology (Netherlands Genomics Initiative). This publication reflects only the authors' views. The European Community is not liable for any use that may be made of the information herein.

Acknowledgements

We would like to thank ServiceXS for performing the real-time PCR experiments using Fluidigm's Dynamic Array on the BioMark instrument. This was partly made possible via the ServiceXS-Fluidigm Application Challenge price, won by Lisa van Baarsen.

Reference List

(1) Baechler EC, Batliwalla FM, Karypis G, Gaffney PM, Ortmann WA, Espe KJ et al. Interferon-inducible gene expression signature in peripheral blood cells of patients with severe lupus. Proc Natl Acad Sci U S A 2003; 100(5):2610-2615.

(2) Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. Curr Opin Immunol 2004; 16(6):801-807.

(3) Ronnblom LE, Alm GV, Oberg KE. Autoimmunity after alpha-interferon therapy for malignant carcinoid tumors. Ann Intern Med 1991; 115(3):178-183.

(4) Feldmann M, Maini RN. Anti-TNF alpha therapy of rheumatoid arthritis: what have we learned? Annu Rev Immunol 2001; 19:163-196.

(5) Elliott MJ, Maini RN, Feldmann M, Kalden JR, Antoni C, Smolen JS et al. Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor alpha (cA2) versus placebo in rheumatoid arthritis. Lancet 1994; 344(8930):1105-1110.

(6) Targan SR, Hanauer SB, van Deventer SJ, Mayer L, Present DH, Braakman T et al. A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group. N Engl J Med 1997; 337(15):1029-1035.

(7) Tracey D, Klareskog L, Sasso EH, Salfeld JG, Tak PP. Tumor necrosis factor antagonist mechanisms of action: a comprehensive review. Pharmacol Ther 2008; 117(2):244-279.

(8) Banchereau J, Pascual V, Palucka AK. Autoimmunity through cytokine-induced dendritic cell activation. Immunity 2004; 20(5):539-550.

(9) Palucka AK, Blanck JP, Bennett L, Pascual V, Banchereau J. Cross-regulation of TNF and IFN-alpha in autoimmune

diseases. Proc Natl Acad Sci U S A 2005; 102(9):3372-3377.

(10) Mavragani CP, Niewold TB, Moutsopoulos NM, Pillemer SR, Wahl SM, Crow MK. Augmented interferon-alpha pathway activation in patients with Sjogren's syndrome treated with etanercept. **Arthritis Rheum 2007; 56(12):3995-4004.**

(11) Aringer M, Graninger WB, Steiner G, Smolen JS. **Safety and efficacy of tumor necrosis factor alpha blockade in systemic** lupus erythematosus: an open-label study. Arthritis Rheum 2004; 50(10):3161-3169.

(12) Rothuizen LE, Buclin T, Spertini F, Trinchard I, Munafo A, Buchwalder PA et al. Influence of interferon beta-1a dose frequency on PBMC cytokine secretion and biological effect markers. J Neuroimmunol 1999; 99(1):131-141.

(13) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. **Ann Rheum Dis 2007; 66(8):1008-1014.**

(14) Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J et al. **A gene expression signature for recent onset rheumatoid** arthritis in peripheral blood mononuclear cells. Ann Rheum Dis 2004; 63(11):1387-1392.

(15) Bokarewa M, Tarkowski A, Lind M, Dahlberg L, Magnusson M. Arthritogenic dsRNA is present in synovial fluid from rheumatoid arthritis patients with an erosive disease course. Eur J Immunol 2008; 38(11):3237-3244.

(16) van Gestel AM, Prevoo ML, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum 1996; 39(1):34-40.

(17) Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. Clin Exp Rheumatol 2005; 23(5 Suppl 39):S93-S99.

(18) van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. Arthritis Rheum 1998; 41(10):1845-1850.

(19) van der Pouw Kraan TC, van Baarsen LG, Rustenburg F, Baltus B, Fero M, Verweij CL. Gene expression profiling in rheumatology. Methods Mol Med 2007; 136:305-327.

(20) Demeter J, Beauheim C, Gollub J, Hernandez-Boussard T, Jin H, Maier D et al. The Stanford Microarray Database: implementation of new analysis tools and open source release of software. Nucleic Acids Res 2007; 35(Database issue): D766-D770.

(21) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. **Genes Immun 2006; 7(6):522-531.**

(22) Bennett L, Palucka AK, Arce E, Cantrell V, Borvak J, Banchereau J et al. **Interferon and granulopoiesis signatures in systemic** lupus erythematosus blood. J Exp Med 2003; 197(6):711-723.

(23) Han GM, Chen SL, Shen N, Ye S, Bao CD, Gu YY. Analysis of gene expression profiles in human systemic lupus erythematosus using oligonucleotide microarray. Genes Immun 2003; 4(3):177-186.

(24) Kirou KA, Lee C, George S, Louca K, Peterson MG, Crow MK. Activation of the interferon-alpha pathway identifies a subgroup of systemic lupus erythematosus patients with distinct serologic features and active disease. Arthritis Rheum 2005; 52(5):1491-1503.

(25) Baechler EC, Bauer JW, Slattery CA, Ortmann WA, Espe KJ, Novitzke J et al. An interferon signature in the peripheral blood of dermatomyositis patients is associated with disease activity. Mol Med 2007; 13(1-2):59-68.

(26) Cantaert T, De RL, Mavragani CP, Wijbrandts CA, Niewold TB, Niers T et al. Exposure to nuclear antigens contributes to the induction of humoral autoimmunity during TNF alpha blockade. Ann Rheum Dis 2008.

(27) Yarilina A, Park-Min KH, Antoniv T, Hu X, Ivashkiv LB. TNF activates an IRF1-dependent autocrine loop leading to sustained expression of chemokines and STAT1-dependent type I interferon-response genes. Nat Immunol 2008; 9(4):378-387.

(28) Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. **Am J Hum Genet 2005;** 76(3):528-537.

(29) Sigurdsson S, Nordmark G, Garnier S, Grundberg E, Kwan T, Nilsson O et al. **A common STAT4 risk haplotype for Systemic**

Lupus Erythematosus is over-expressed, correlates with anti-dsDNA production and shows additive effects with two IRF5 risk alleles. Hum Mol Genet 2008.

(30) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007; 357(10):977-986.

(31) Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, de la Serna AR, Fernandez-Gutierrez B, Castaneda S et al. Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. Arthritis Rheum 2008; 58(5):1264-1274.

(32) Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Wiman AC, Alfredsson L et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 2007; 56(7):2202-2210.

Table S1. Real-time PCR assay information

Primer sequences for regular quantitative realtime PCR using SybrGreen:

Symbol	product size (bp)	Forward primer sequence	Reverse primer sequence
18SRNA	84	CCGAGTAAGTGCGGGTCATAA	CCATCCAATCGGTAGTAGCG
RSAD2	90	GTGGTTCCAGAATTATGGTGAGTATTT	CCACGGCCAATAAGGACATT
OAS1	175	TGCGCTCAGCTTCGTACTGA	GGTGGAGAACTCGCCCTCTT
IFI44L	77	CCGAGCGGTATAGGATATATTCTGTT	TGCTCCTTCTGCCCCATCTA

Pre-designed Taqman Gene Expression Assays used for Fluidigm's BioMark™ Real-Time PCR System :

Symbol	Assay ID	Gene Name
IFI6	Hs00242571_m1	interferon, alpha-inducible protein 6
IFIT1	Hs00356631_g1	interferon-induced protein with tetratricopeptide repeats 1
IFITM1	Hs00705137_s1	interferon induced transmembrane protein 1 (9-27)
IRF2	Hs01082884_m1	interferon regulatory factor 2
ISG15	Hs01921425_s1	ISG15 ubiquitin-like modifier
OAS2	Hs00942643_m1	2'-5'-oligoadenylate synthetase 2, 69/71kDa
RSAD2	Hs00369813_m1	radical S-adenosyl methionine domain containing 2
SAMD9L	Hs00541567_s1	sterile alpha motif domain containing 9-like
SERPING1	Hs00934329_m1	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1,(angioedema, hereditary)
EPSTI1	Hs00264424_m1	epithelial stromal interaction 1 (breast)
IFI35	Hs00413458_m1	interferon-induced protein 35
IFI44L	Hs00199115_m1	interferon-induced protein 44-like
Mx2	Hs00159418_m1	myxovirus (influenza virus) resistance 2 (mouse)
OAS1	Hs00242943_m1	2',5'-oligoadenylate synthetase 1, 40/46kDa
18SRNA	Hs99999901_s1	Eukaryotic 18S rRNA;18S ribosomal RNA
LGALS3BP	Hs01003086_m1	lectin, galactoside-binding, soluble, 3 binding protein

Chapter 4.3

Pharmacogenomics of Interferon-ß Therapy in Multiple Sclerosis: Baseline IFN Signature Determines Pharmacological Differences between Patients

4

Lisa G. M. van Baarsen[1,2*], Saskia Vosslamber[1,2*], Marianne Tijssen[2], Josefien M. C. Baggen[1], Laura F. van der Voort[3], Joep Killestein[3], Tineke C. T. M van der Pouw Kraan[1], Chris H. Polman[3], Cornelis L. Verweij[1,2]

[1]Department of Molecular Cell Biology and Immunology, [2]Department of Pathology, [3]Department of Neurology, VU Medical Center, Amsterdam, The Netherlands

** These authors contributed equally to this work*

PLoS ONE 2008 Apr 2;3(4):e1927

Abstract

Background: Multiple sclerosis (MS) is a heterogeneous disease. In order to understand the partial responsiveness to IFNß in Relapsing Remitting MS (RRMS) we studied the pharmacological effects of IFNß therapy.

Methodology: Large scale gene expression profiling was performed on peripheral blood of 16 RRMS patients at baseline and one month after the start of IFNß therapy. Differential gene expression was analyzed by Significance Analysis of Microarrays. Subsequent expression analyses on specific genes were performed after three and six months of treatment. Peripheral blood mononuclear cells (PBMC) were isolated and stimulated *in vitro* with IFNß. Genes of interest were measured and validated by quantitative realtime PCR. An independent group of 30 RRMS patients was used for validation.

Principal Findings: Pharmacogenomics revealed a marked variation in the pharmacological response to IFNß between patients. A total of 126 genes were upregulated in a subset of patients whereas in other patients these genes were downregulated or unchanged after one month of IFNß therapy. Most interestingly, we observed that the extent of the pharmacological response correlates negatively with the baseline expression of a specific set of 15 IFN response genes ($R = -0.7208$; $p = 0.0016$). The negative correlation was maintained after three ($R= -0.7363$; $p= 0.0027$) and six ($R= -0.8154$; $p= 0.0004$) months of treatment, as determined by gene expression levels of the most significant correlating gene. Similar results were obtained in an independent group of patients ($n=30$; $R = -0.4719$; $p = 0.0085$). Moreover, the *ex vivo* results could be confirmed by *in vitro* stimulation of purified PBMCs at baseline with IFNß indicating that differential responsiveness to IFNß is an intrinsic feature of peripheral blood cells at baseline.

Conclusion: These data imply that the expression levels of IFN response genes in the peripheral blood of MS patients prior to treatment could serve a role as biomarker for the differential clinical response to IFNß.

4

Introduction

Multiple sclerosis (MS) is a common inflammatory disease of the central nervous system characterized by progressive neurological dysfunction. The disease has a heterogeneous nature, which is reflected in the clinical presentation, ranging from mild to severe demyelinating disease. No curative therapy is currently available, and the majority of affected individuals are ultimately disabled.[1]

IFNs were the first agents to show clinical efficacy in RRMS. Interferon beta (IFNß) decreases clinical relapses, reduces brain disease activity, and possibly slows down progression of disability. However, therapy is associated with a number of adverse reactions, including flu-like symptoms and transient laboratory abnormalities. Moreover, the response to IFNß is partial, i.e. disease activity is suppressed by only

about one third.[2] Clinical experience suggests that there are IFN'responders' as well as 'non responders', however clear criteria for such classification are still lacking.[3] In the absence of predictive biomarkers the question remains who will respond to therapy and who to treat when inconvenience and costs are significant.
Part of the unresponsiveness to IFNß can be explained by immunogenicity. However, since not all unresponsive patients develop neutralizing antibodies (Nabs), and Nabs can disappear over time,[4–7] other mechanisms have to be involved to explain unresponsiveness. Hence, there have to be biological disease mechanisms in a subpopulation of patients that results in insensitivity or resistance to the effects of IFNs. This implies that pharmacological responses may differ between patients, leading to inter-individual differences in clinical efficacy. We hypothesize that an in depth understanding of the pharmacological factors underlying the therapeutic mechanisms and therapy unresponsiveness is the key for the identification of predictive markers.
In normal physiology type I IFNs achieve their biological effects by binding to multi-subunit receptors IFNAR-1 and -2 on the cell surface, thereby initiating a complex cascade of intracellular secondary messengers that emerge in two divergent pathways. One pathway, leads to activation of the transcription factor IFN-stimulated gene factor 3 (ISGF3), a complex of phosphorylated Signal Transducer and Activator of Transcription (STAT) 2 with STAT1 and IFN regulatory factor 9 (IRF-9; p48) that binds to the IFN-stimulated response element (ISRE) present in multiple genes.[8,9] The other pathway involves STAT2/1 and STAT2/3 heterodimers and STAT1 homodimer (IFN-α-activated factor, AAF), which bind to the IFN gamma-activated sequence (GAS) response element.[9–12] Ultimately, the IFN-induced activation of ISRE and GAS enhancer elements switch on a wide variety of genes[13] leading to specific transcriptional changes.

With the aid of genomics technology, we are now in a position to provide sufficient knowledge to determine pharmacological outcomes that will allow us to search for predictors of therapeutic outcomes. Previously we demonstrated that gene expression signatures in MS may differ significantly between patients.[14] We found that a subgroup of MS was characterized by an increased expression of an immune defense response gene set, including a type I IFN response signature. Here we generated and analyzed pre- and post- IFNß treatment gene expression patterns of RRMS patients with the aim of identifying pre-existing and/or drug-induced signatures that will allow us to make predictions on the expected pharmacological effects of IFNß treatment. We show that the expression level of IFN response genes prior to treatment, could serve a role as biomarker for the pharmacological differences between patients with MS at the molecular level.

Methods

Patients

A first group of 16 Dutch patients (10 females and 6 males) and a second group of 30 Dutch patients (17 females and 13 males) with clinically definite relapsing-remitting MS was recruited from the outpatient clinic of the MS Centre Amsterdam. Mean age at start of IFNß therapy for the test group is 40.6 ± 7.7, mean EDSS is 2.3 ± 1.3 (range 1 - 6). Blood samples were obtained at a fixed time during the day just before treatment and 1, 3 and 6 months after start of the therapy. Patients received either Avonex (n=4), Betaferon (n=7), Rebif 22 (=2) or Rebif 44 (n=3). For the validation group, mean age at start of IFNß therapy is 34.0 ± 9.9, mean EDSS 2.3 ± 1.1 (range 0 - 4.5). Patients received either Avonex (n=7), Betaferon (n=8), Rebif 22 (n=4) or Rebif 44 (n=11).

The study was approved by the ethics committee of the VUmc and all patients provided written informed consent.

Blood sampling

From each patient blood was drawn into one PAXgene tube (PreAnalytix, GmbH, Germany) and three heparin tubes (Beckton Dickinson, Alphen a/d Rijn, Netherlands). After blood collection, tubes were transferred from the clinic to the lab within one hour in order to isolate fresh peripheral blood mononuclear cells (PBMCs) from heparinized blood using lymphoprep (Axis-Shield, Lucron) density gradient centrifugation. PAXgene tubes were stored at room temperature (RT) for two hours to ensure complete lyses of all blood cells after which tubes were stored at –20 until RNA isolation. Total RNA was isolated within 7 months after storage. Tubes were thawed 2 hours at RT prior to RNA isolation. Next, RNA was isolated using the PreAnalytix RNA isolation kit according to the manufacturers' instructions, including a DNAse (Qiagen, Venlo, Netherlands) step to remove genomic DNA. Quantity and purity of the RNA was tested using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, Delaware USA).

Microarray hybridization

We used 43K cDNA microarrays from the Stanford Functional Genomics Facility (http://microarray.org/sfgf/) printed on aminosilane-coated slides containing ~20.000 unique genes. First DNA spots were UV-crosslinked to the slide using 150-300 mJoules. Prior to sample hybridisation, slides were prehybridized at 42 degrees Celsius for 15 minutes in a solution containing 40% ultra pure formamide (Invitrogen, Breda, Netherlands), 5% SSC (Biochemika, Sigma), 0.1% SDS (Fluka Chemie, GmbH, Switserland) and 50 µg/ml BSA (Panvera, Madison, USA). After prehybridization slides were briefly rinsed in MilliQ water, thoroughly washed in boiling water and 95% ethanol and air-dried. Sample preparation and microarray hybridisation was performed as described previously,[15] apart from the different postprocessing and prehybridization described above.

Microarray analysis

Data storage and filtering was performed using the Stanford Microarray Database[16] (http://genome-www5.stanford.edu//) as described previously.[14] Statistical Analysis of Microarrays [17] (SAM) was used to determine significantly differential expressed genes. A gene was considered as significantly differential expressed if the False Discovery Rate (FDR) was equal to or less than 5%. Cluster analysis[18] was used to define clusters of co-coordinately changed genes after which the data was visualized using Treeview.

Microarray data in this paper are stored in the publicly accessible Stanford Microarray Database website http://smd.stanford.edu/ which supports the MIAME guidelines. In addition, data is stored in the Gene Expression Omnibus (GSE10655). The National Center for Biotechnology Information (www.ncbi.nlm.nih.gov/) Genbank accession numbers for the genes and gene products discussed in this paper are listed in Table S1.

Realtime PCR

RNA (0.5 μg) was reverse transcribed into cDNA using a Revertaid H-minus cDNA synthesis kit (MBI Fermentas, St. Leon-Rot, Germany) according to the manufacturers' instructions. Quantitative realtime PCR was performed using an ABI Prism 7900HT Sequence detection system (Applied Biosystems, Foster City, CA, USA) using SybrGreen (Applied Biosystems). Primers were designed using Primer Express software and guidelines (Applied Biosystems) and are listed in table 1. To calculate arbitrary values of mRNA levels and to correct for differences in primer efficiencies a standard curve was constructed. Expression levels of target genes were expressed relative to housekeeping gene *glyceraldehydes-3-phosphate dehydrogenase* (*GAPDH*).

Table 1. Primers used for quantitative realtime PCR

Genes	Genbank accession nr.	Sense primer	Antisense primer	Length PCR product (bp)
MxA	NM 002462	TTCAGCACCTGATGGCCTATC	GTACGTCTGGAGCATGAAGAACTG	92
OAS1	NM 016816	TGCGCTCAGCTTCGTACTGA	GGTGGAGAACTCGCCCTCTT	175
STAT1	NM 007315	TGCATCATGGGCTTCATCAGC	GAAGTCAGGTTCGCCTCCGTTC	156
RSAD2	NM 080657	GTGGTTCCAGAATTATGGTGAGTATTT	CCACGGCCAATAAGGACATT	90
IRF7	NM 004031	GCTCCCCACGCTATACCATCTAC	GCCAGGGTTCCAGCTTCAC	99
ISG15	NM 005101	TTTGCCAGTACAGGAGCTTGTG	GGGTGATCTGCGCCTTCA	151
IFNß	NM 002176	ACAGACTTACAGGTTACCTCCGAAAC	CTCCTAGCCTGTCCCTCTGGGACTGG	93

In vitro study

Freshly isolated PBMCs were washed using PBS containing 1% fetal calf serum (FCS; BioWhittaker, Cambrex) and plated in 24-wells culture plates at a density of $2x10^6$ cells per ml per well. Cells were left unstimulated or activated with 10 Units recombinant IFNß (Abcam, Cambridge, UK) for 4h after which RNA was isolated

using the Rneasy Qiagen RNA isolation kit (Qiagen) according to the manufacturers' instructions. A DNAse (Qiagen) step was included to remove genomic DNA. Quantity and purity of the RNA was tested using the Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, Delaware USA)

Statistical analysis

Correlation analyses were performed using Graphpad Prism 4 software. First, data was tested for normal distribution. For normally distributed data, a Pearson correlation was used. A Spearman rank correlation was calculated in case of nonparametric distribution of the data. Correlations were considered significant if p-values were less than 0.05.

Results

Pharmacogenomics of IFNß therapy in MS

In order to understand the pharmacological effects of IFNß therapy we analysed the peripheral blood gene expression profiles of 16 RRMS patients at baseline and one month after the start of therapy. Two class paired analysis using Significant Analysis of Microarrays (SAM) at a False Discovery Rate (FDR) of less than 5% between pre- and post-therapy data was applied to identify genes that significantly changed in expression after IFNß treatment. Surprisingly, only 3 genes, "Interferon

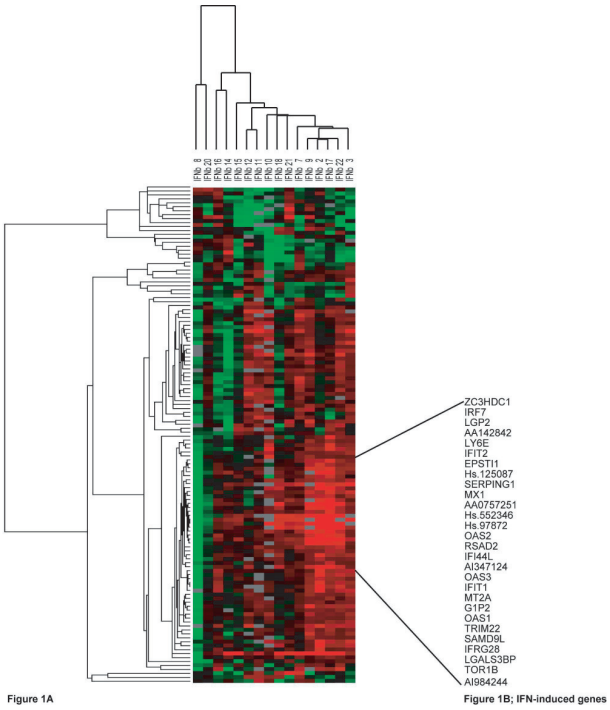


Figure 1

A. Biological response to IFNß therapy in MS patients.
Two-way hierarchical cluster analyses using gene expression ratio's (biological response). This diagram contains genes that were at least two-fold up- or downregulated after IFNß therapy in at least seven patients. Upregulated genes after therapy are indicated by a red colour, downregulated by a green colour and genes that show no differences in expression after therapy are indicated in black.

B. Cluster of IFN-induced genes.
Selection of genes clustering together based on similar biological response profiles within the patient group. The genes clustered together with a correlation of 0.925 and are known to be induced by IFN. The mean expression ratio of all genes in this IFN cluster is referred to as the biological IFN response.

alpha-inducible protein 27" (IFI27), "Tripartite motif-containing 69" (TRIM69) and "Epithelial stromal interaction protein 1 (breast)" (EPSTI1), showed a significant change.

Given the heterogeneous nature of MS we questioned whether the observed poor yield of response genes upon IFNß treatment of the whole MS cohort could be a reflection of averaging out differences as a consequence of variation in pharmacological responsiveness between the patients. To test this hypothesis we investigated the pharmacological response at the individual patient level by calculating for each patient and for each gene the ratio of gene expression pre- vs. post therapy (log-2 ratios). We selected genes that showed at least a two-fold change in expression after IFNß treatment in at least 7 patients. A total of 126 genes met this criteria and were subsequently subjected to a two-way hierarchical (unsupervised) cluster analysis (Figure 1A). Compliant with our hypothesis, this analysis showed a marked variation in biological response to IFNß between patients. Some patients showed upregulated genes, whereas in other patients the same genes were downregulated or unchanged after IFNß therapy. As anticipated, part of this gene expression pattern is consistent with expression of known IFN response genes [13]. We next selected the cluster of genes showing the most inter-individual variation resulting in 28 IFN-induced genes (Table S1) that clustered tightly together (R=0.925) indicating a coordinate regulation of these genes (Figure 1B). The expression data of some of the IFN-induced genes was validated by quantitative realtime PCR and showed a good correlation with the microarray data (Table 2). These findings confirmed the hypothesis that there exists considerable variation in the pharmacological effects of IFNß between patients with RRMS.

Table 2. Correlation between microarray data and realtime PCR data

Genes	p value	R value
MxA	0.0188	0.4335
OAS1	<0.0001	0.6972
STAT1	<0.0001	0.7371
RSAD2	<0.0001	0.7086
IRF7	0.0014	0.5648
ISG15	<0.0001	0.7051

Relationship between pharmacological response and baseline gene expression levels

Previously, we demonstrated significant differences in the expression of type I IFN-induced genes between untreated RRMS patients.[14] Here we investigated whether there is a relationship between the differential in vivo responsiveness to IFNß and baseline expression levels of IFN-induced genes. Therefore, we tested for each patient whether there is an association between the mean expression levels of the IFN response gene cluster (shown in Figure 1B) before therapy with the response ratio after therapy. This analysis demonstrated that the mean baseline expression of the 28 IFN response genes negatively correlates with the in vivo IFN-

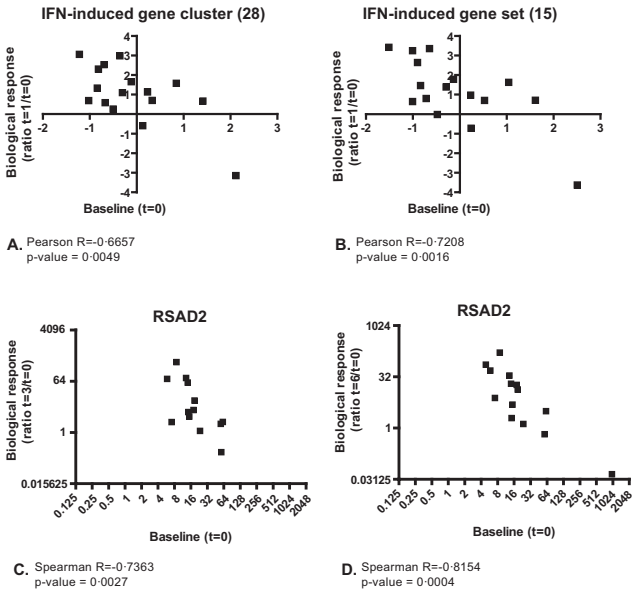


Figure 2. Correlation between baseline and biological response to IFNß therapy.
Biological responses were calculated, using a set of IFN-induced genes (A and B) or a single IFN-induced gene (C and D) and correlated with baseline levels, resulting in a significant negative correlation. In C and D the expression levels of RSAD2 is measured by quantitative realtime PCR and normalized to the expression levels of GAPDH. A. IFN cluster as described in Figure 1B; B. Selection of 15 genes; C. Biological response after three months, using RSAD2 gene expression levels; D Biological response after six months using RSAD2 gene expression levels.

induced response levels (p = 0.0049 and R = -0.6657) (Figure 2A), suggesting that the baseline gene expression level of these genes could serve a role as predictive marker for the pharmacological responsiveness to IFNß.

In order to create a gene set that best predicts the pharmacological response to IFNß we selected those genes whose expression shows the most significant negative correlation between baseline and biological response (with a cut off of p<0.01 and R <-0.65). This resulted in a gene set containing 15 genes (Table 3). Comparing baseline gene expression levels and biological response using the average of these 15 genes revealed a significant negative correlation (R = -0.7208; p = 0.0016) (figure 2B). To exclude a potential bias of the gene selection at baseline, we analyzed the correlation of the biological response determined by the mean expression value of the selected 15 IFN-induced genes with the baseline values of all genes on the array. This resulted in three additional genes (IFI44L, MT1E and IMAGE:1879725; R <-0.65 and variance >1.00) that significantly correlated with the pharmacological response to IFNß therapy. Although these genes did not cluster tightly together with the previously selected genes, they may be important in the response to IFNß.

To investigate whether the observed negative correlation between baseline and treatment induced changes are stable over time we measure the expression level of the most significant correlating gene (RSAD2; see Table 3) again after three and six months of IFNß therapy. The negative correlation between baseline expression level and biological response was maintained after 3 months (p = 0.0027, R =-0.7363) and 6 months (p = 0.0004, R = -0.8154) of therapy (figure 2C and D). To validate our results, we measured expression levels of RSAD2 in a second independent group of

patients (n=30) before and after IFNß treatment. In this independent study group we confirmed the negative correlation between baseline gene expression level and treatment induced biological response (p<0.0085 and R=-0.4719).

Table 3. Correlation between baseline and therapy induced (ratio) expression levels measured at single gene level

Symbol	Accession number	p value	R value
RSAD2	NM_080657	0.0011	-0.7983
IFIT1	NM_001548	0.0004	-0.7746
MX1	NM_002462	0.0006	-0.7619
ISG15	NM_005101	0.0008	-0.7532
IMAGE:1926927	AI347124	0.0026	-0.7168
EPSTI1	NM_001002264	0.0059	-0.7162
Transcribed locus	Hs.552346	0.0038	-0.6977
IRF7	NM_004031	0.0029	-0.6925
IMAGE:545138	5'EST AA075776; 3'EST AA075725	0.0065	-0.6881
LY6E	NM_002346	0.0035	-0.6834
OAS1	NM_016816	0.0051	-0.6822
OAS3	NM_006187	0.0076	-0.6787
IMAGE:504372	AA142842	0.0087	-0.6707
SERPING1	NM_000062	0.0064	-0.6688
Transcribed locus	Hs.97872	0.0047	-0.6677
IFI44L	NM_006820	0.0196	-0.6353
Transcribed locus	Hs.125087	0.0108	-0.6175
MT2A	NM_005953	0.011	-0.6166
TRIM22	NM_006074	0.0118	-0.6115
SAMD9L	NM_152703	0.0121	-0.6102
IMAGE:2562181=OAS2	NM_002535	0.0203	-0.591
OAS2	NM_002535	0.0169	-0.5865
DHX58	NM_024119	0.0222	-0.5842
PARP12	NM_022750	0.0343	-0.5482
TOR1B	NM_014506	0.0532	-0.4914
IFIT2	NM_001547	0.086	-0.4426
RTP4	NM_022147	0.1759	-0.369
LGALS3BP	NM_005567	0.314	-0.279

Comparative analyses of different treatment regimens

Since in the present study different pharmaceutical IFNβ preparations were used for treatment, we wanted to exclude the possibility of potential differences in pharmacokinetics and exposure as an explanation for our findings. Different studies have indicated no or negligible differences in bioavailability between different treatment preparations and routes of administration [19,20]. To exclude a possible bias in our results due to differences in frequency of injection [20,21] we

divided our patients in two groups based on frequency of injection and compared their biological responses. One group of patients (group A) consists of patients with weekly treatment (Avonex) and the other group of patients (group B) who are treated three to four times a week (Rebif and Betaferon). Comparison of the response rates between the treatment groups revealed a similar range of response levels independent of the treatment regimen for both the test cohort (group A, n=4 and group B, n=12) based on microarray data, and the validation cohort (group A n=6 and group B n=24) based on quantitative PCR data (Figure 3). To provide further evidence that our results were not influenced by the frequency of injection we confirmed the negative correlation between the response rate and baseline IFN response gene expression in the group of frequently dosed patients (group B: test cohort (n=12), R=-0.8361, p=0.0007; validation cohort (n=24), R=-0.4513, p=0.0269).

Altogether, these results reveal that the observed negative correlation between baseline IFN signature and the extent of the biological response is not biased by the treatment regimen.

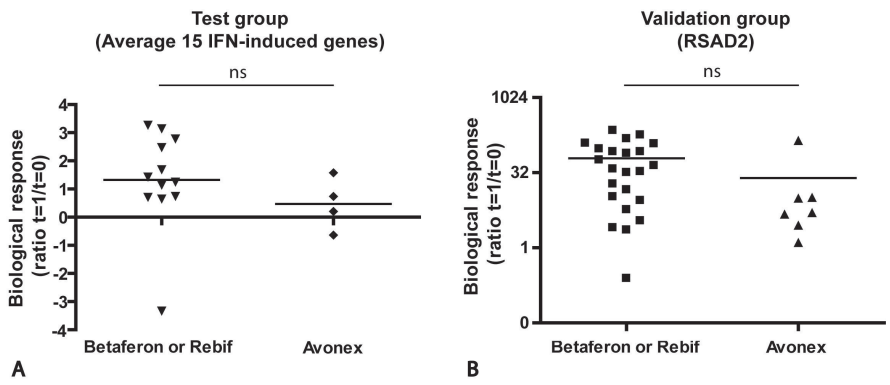


Figure 3. Comparative analysis between different treatment regimens.
Comparison of biological response of Avonex treated patients and Betaferon or Rebif treated patients. A. Average biological response using the set of 15 IFN-induced genes in the test group of 16 RRMS patients; B. Biological response using PCR based gene expression levels for RSAD2 in the second independent validation group of 30 RRMS patients.

Confirmation of ex vivo findings by in vitro IFNß stimulation of PBMC isolated at baseline

To further confirm that the observed inter-individual pharmacological differences were a consequence of differential responsiveness of peripheral blood cells and to exclude i. blood sampling error differences because of possible differential time-intervals between blood sampling and injection of IFNß, and ii. interference of inhibitory plasma proteins such as neutralizing antibodies, we performed an *in vitro* cell stimulation assay. Therefore we used purified PBMCs isolated prior to treatment, which were cultured for 4 hours in the presence of recombinant IFNß. To analyze the *in vitro* response to IFNß at baseline we measured the expression

of a selected set of three known IFNß response genes and IFNß itself in resting and IFNß treated PBMCs by quantitative realtime PCR. The selected IFNß response genes were i. *RSAD2*, which showed the most significant correlation of biological response versus baseline at single gene level (Table 3), ii. *MxA*, which showed a good negative correlation and is known as a marker of IFN bioactivity, [22] and iii. *STAT1*, which is one of the components important for IFNß signaling. We hypothesized that baseline expression level of these genes influences subsequent IFNß signaling upon treatment. We compared the *in vitro* biological response of these genes to the mean *in vivo* biological response of the selected 15 genes. For all genes a significant correlation was revealed between the *in vitro* and *in vivo* biological response (Table 4). From these results we concluded that the differential IFNß responsiveness in MS is a consequence of intrinsic differences of peripheral blood cells in their responsiveness to IFNß. Moreover, the consistency between the *in vivo* and *in vitro* response to IFNß provides further evidence to exclude the involvement of different types and dosages of treatment on the observed pharmacological differences.

Table 4. Correlation between biological responses of single IFN-induced genes measured in vitro and mean biological response (using 15 genes) measured in vivo

Genes	p value	R value
RSAD2	0.0012	0.7518
MxA	0.0280	0.6064
STAT1	0.0100	0.6614
IFNß	0.0036	0.7675

Biological IFN response and clinical parameters

The results described above could point towards a method to predict responsiveness to IFNß therapy based on baseline expression levels of IFN-induced genes. In the clinic the response status of a patient is measured by evaluation of Expanded Disability Status Scale (EDSS) progression, relapse rate and disease activity on Magnetic Resonance Imaging (MRI). For the first patient group (n=16) EDSS progression, number of steroid interventions and relapse rate two years before initiation of treatment were assessed retrospectively and compared to the first two years after start of treatment. With this limited set of response criteria no association with the predictive pharmacological gene set of 15 IFN induced genes could be observed.

▪ Discussion

Our results reveal that RRMS patients show a heterogeneous pharmacological response to IFNß therapy. In some patients we demonstrate that administered exogenous IFNß induces functional activation of the IFN pathway, whereas other patients do not reveal a functional IFNß response. The latter are characterized by a biomarker profile reflecting a saturated IFN activation pathway prior to treatment.

Hence the baseline expression of the biomarker profile reflecting the baseline status of the IFN activity negatively correlates with the pharmacological effects of IFNß treatment. This indicates that the baseline expression levels of the selected set of 15 IFN-induced genes can be used as a predictive marker for the responsiveness to IFNß treatment.
Thus patients with clinically defined similar disease may have intrinsic different modes of immune status. These findings make more evident the complexity of the disease and the relationship to therapy responsiveness.

Although different regimens of IFNß treatment were used in this study evidence is available that this does not affect our conclusions.
Firstly, there is accumulating evidence that there is no or little difference between different types of IFNβ in terms of their biological activity and routes of administration [19,20]. Extent and duration of clinical and biologic effects were independent of the route of administration of IFNβ. Rebif when given *s.c.* or *i.m.* was found to be bioequivalent to Avonex [23,24]. Moreover, there were no major differences between the results with IFNβ1a and 1b in the duration of the changes in the pharmacodynamic markers after the two routes of injection [25,26].
Secondly, we excluded a possible bias in our results due to frequency of injection by analyzing different treatment groups separately. No significant differences in the range of biological response levels between Avonex treated patients and Rebif or Betaferon treated patients were observed, and selection of the high-frequently (Rebif and Betaferon) dosed patients by excluding weekly–treated (Avonex) patients from our analyses still resulted in a negative correlation between baseline IFN levels and biological response rate.
Thirdly, in the present study we show that the observed negative correlation between biological response and baseline levels of IFN induced genes is consistently observed over time, at one, three and six months after start of the therapy.
Finally, we showed that response-rates of *in vitro* stimulated PBMC isolated prior to treatment are consistent with those of the *ex vivo* results. These results convincingly supported the conclusion that the *in vivo* biological response is independent of differences in treatment regimens and interfering serum proteins such as neutralizing antibodies (Nabs).
Hence, we concluded that the inter-individual variation in pharmacological response to IFNß therapy is an intrinsic property of the peripheral blood cell compartment.

Several investigators have recently reported on transcription based responses to IFNß in MS. Baranzini and colleagues [27] used a pre-selected set of 70 genes and reported that (un)supervised two-way hierarchical clustering does not reveal significantly differential expressed genes between responders and non-responders. Using quadratic discriminant analysis-based integrated Bayesian inference system

they found a gene triplet consisting of apoptosis-related genes as best predictive for good responder versus poor responder classification. Most of the 70 genes they selected are represented on our microarray but we didn't observe a difference for these genes using a gene-by-gene approach. However, the majority of genes that we found as predictive for responsiveness using an open survey approach were not present in the gene set selected by Baranzini and colleagues and therefore not identified in their study. A careful comparison between the different IFNβ pharmacogenomics studies [28,29] learns that there is consistency between these reports and our data with respect to the heterogeneity of the IFNβ response. Although not explicitly mentioned in these reports, we learned that they contained evidence for inter-individual differences in response to IFNβ. Overall, despite basic differences in the designs, we confirm and extend the trends observed in these reports with respect to the heterogeneity in treatment response rates. In addition, our paired analysis method provides an ideal approach for a patient centric mode of data analysis and discloses significant differences in the expression of an IFN driven response gene set at baseline in relation to the pharmacological response. Our findings provide a perfect explanation for the inter-individual variation in the pharmacological responses mentioned above.

Our data based on paired analysis at the individual patient level clearly show that there is evidence for differences in IFNß responsiveness between patients with MS. The inter-individual differences in IFNß responsiveness may be the result of genetic variation in the IFNß biology.[2,30] Feng and colleagues [31] showed that IFN-induced levels of mRNA and protein for IFN-regulatory genes (*IRF-1* and *IRF-2*) and antiviral genes (*MxA* and *2´, 5´-OAS*) were significantly lower in PBMC from patients with clinically active MS compared to normal controls. They demonstrated that clinical disease activity was related to decreased phosphorylation of Ser-STAT-1 and proposed that this could be a mechanism explaining a defective IFN response. Whereas these studies provided insight into the IFN responsiveness in terms of a group average the issue of inter-individual heterogeneity was not addressed. Other mechanisms that could account for differential responsiveness to IFNß include variation in activity of inhibitory transcription factors. Evidence exists that crosstalk with other cytokine-activated pathways, could cause tachyphylaxis to type I IFNs. Although type I IFNs have an essential function in mediating innate immune responses against viruses, they may already be produced at very low levels in the absence of viral infections [32] in serum of a subset of MS patients. Since e.g. IFNα is known to desensitize further responses to IFNs, the presence of low endogenous IFNs could block IFNß-induced signals.[33,34]

This explorative pilot study suggests a predictive value of baseline gene expression levels of IFN-induced genes. Since the molecular differences most likely reflect distinct pathophysiologic processes underlying disease, it is tempting to speculate

that these differences will predict individual responsiveness to treatment. Clinical response to IFNß may be determined by disability progression and relapse rate. Because MS is a chronic disease with an unpredictable clinical course it remains difficult to assess clinical responder status at an individual patient level. A more objective method for determining disease activity is the measurement of MRI parameters, e.g. CNS atrophy measures or T1 gadolinium enhancing or the appearance of new T2 lesions.[3,35,36] However, using these methods it is still extremely difficult to precisely define the state of responsiveness after a short period of treatment or preferably before start of the treatment. These facts emphasize the importance of finding pharmacological predictors and/or determinants for treatment responsiveness. We realize that the design of this study does not allow any firm conclusions to be drawn concerning the clinical parameters associated with the molecular phenotype.

Hence, further studies in a large cohort of patients starting IFNß treatment are needed to validate and further investigate the predictive value of baseline IFN response gene expression levels and it is of great importance to find a correlation between clinical parameters and the biological IFN response. In future, molecular stratification of patients at baseline may be helpful in assembling homogeneous populations of patients, which will improve the likelihood of observing drug efficacy in clinical trials.

Acknowledgments

The authors gratefully acknowledge the staff and patients of the department of Neurology at the VUMC hospital who participated in this study. In addition we would like to thank Lisa van Winsen who at first recruited the patients for the start of this study.

References

1. Hafler DA, Slavik JM, Anderson DE, O'Connor KC, De JP, et al. (2005) Multiple sclerosis. Immunol Rev 204: 208-231.
2. The IFNB Multiple Sclerosis Study Group and The University of British Columbia MS/MRI Analysis Group. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. (1995)Neurology 45: 1277-1285.
3. Rudick RA, Lee JC, Simon J, Ransohoff RM, Fisher E (2004) Defining interferon beta response status in multiple sclerosis patients. Ann Neurol 56: 548-555.
4. Reske D, Walser A, Haupt WF, Petereit HF (2004) Long-term persisting interferon beta-1b neutralizing antibodies after discontinuation of treatment. Acta Neurol Scand 109: 66-70.
5. Rice GP, Paszner B, Oger J, Lesaux J, Paty D, et al. (1999) The evolution of neutralizing antibodies in multiple sclerosis patients treated with interferon beta-1b. Neurology 52: 1277-1279.
6. Sibley WA, Bamford CR, Clark K (1985) Clinical viral infections and multiple sclerosis. Lancet 1: 1313-1315.
7. Vartanian TK, Zamvil SS, Fox E, Sorensen PS (2004) Neutralizing antibodies to disease-modifying agents in the treatment of multiple sclerosis. Neurology 63: S42-S49.

8. Ghislain JJ, Fish EN (1996) Application of genomic DNA affinity chromatography identifies multiple interferon-alpha-regulated Stat2 complexes. J Biol Chem 271: 12408-12413.

9. Li X, Leung S, Qureshi S, Darnell JE, Jr., Stark GR (1996) Formation of STAT1-STAT2 heterodimers and their role in the activation of IRF-1 gene transcription by interferon-alpha. J Biol Chem 271: 5790-5794.

10. Brierley MM, Fish EN (2005) Functional relevance of the conserved DNA-binding domain of STAT2. J Biol Chem 280: 13029-13036.

11. Ghislain JJ, Wong T, Nguyen M, Fish EN (2001) The interferon-inducible Stat2:Stat1 heterodimer preferentially binds in vitro to a consensus element found in the promoters of a subset of interferon-stimulated genes. J Interferon Cytokine Res 21: 379-388.

12. Takaoka A, Taniguchi T (2003) New aspects of IFN-alpha/beta signalling in immunity, oncogenesis and bone metabolism. Cancer Sci 94: 405-411.

13. Der SD, Zhou A, Williams BR, Silverman RH (1998) Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. Proc Natl Acad Sci U S A 95: 15623-15628.

14. van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, et al. (2006) A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 7: 522-531.

15. van der Pouw Kraan TC, Baarsen EGM, Rustenburg F, Baltus B, Fero M, et al. (2007) Gene expression profiling in rheumatology. In: Cope AP, editors. Arthritis Research. The Humana Press Inc. pp.

16. Ball CA, Awad IA, Demeter J, Gollub J, Hebert JM, et al. (2005) The Stanford Microarray Database accommodates additional microarray platforms and data formats. Nucleic Acids Res 33: D580-D582.

17. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98: 5116-5121.

18. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863-14868.

19. Gilli F, Marnetto F, Caldano M, Sala A, Malucchi S, et al. (2005) Biological responsiveness to first injections of interferon-beta in patients with multiple sclerosis. J Neuroimmunol 158: 195-203.

20. Markowitz CE (2007) Interferon-beta: mechanism of action and dosing issues. Neurology 68: S8-11.

21. Revel M (2003) Interferon-beta in the treatment of relapsing-remitting multiple sclerosis. Pharmacol Ther 100: 49-62.

22. Kracke A, von WP, Al-Masri AN, Dalley G, Windhagen A, et al. (2000) Mx proteins in blood leukocytes for monitoring interferon beta-1b therapy in patients with MS. Neurology 54: 193-199.

23. Alam J, Goelz S, Rioux P, Scaramucci J, Jones W, et al. (1997) Comparative pharmacokinetics and pharmacodynamics of two recombinant human interferon beta-1a (IFN beta-1 a) products administered intramuscularly in healthy male and female volunteers. Pharm Res 14: 546-549.

24. Ravandi F, Estrov Z, Kurzrock R, Breitmeyer JB, Maschek BJ, et al. (1999) A phase I study of recombinant interferon-beta in patients with advanced malignant disease. Clin Cancer Res 5: 3990-3998.

25. Salmon P, Le Cotonnec JY, Galazka A, bdul-Ahad A, Darragh A (1996) Pharmacokinetics and pharmacodynamics of recombinant human interferon-beta in healthy male volunteers. J Interferon Cytokine Res 16: 759-764.

26. Sturzebecher S, Maibauer R, Heuner A, Beckmann K, Aufdembrinke B (1999) Pharmacodynamic comparison of single doses of IFN-beta1a and IFN-beta1b in healthy volunteers. J Interferon Cytokine Res 19: 1257-1264.

27. Baranzini SE, Mousavi P, Rio J, Caillier SJ, Stillman A, et al. (2005) Transcription-based prediction of response to IFNbeta using supervised computational methods. PLoS Biol 3: e2.

28. Weinstock-Guttman B, Badgett D, Patrick K, Hartrich L, Santos R, et al. (2003) Genomic effects of IFN-beta in multiple sclerosis patients. J Immunol 171: 2694-2702.

29. Sturzebecher S, Wandinger KP, Rosenwald A, Sathyamoorthy M, Tzou A, et al. (2003) Expression profiling identifies responder and non-responder phenotypes to interferon-beta in multiple sclerosis. Brain 126: 1419-1429.

30. Zhuang H, Kosboth M, Lee P, Rice A, Driscoll DJ, et al. (2006) Lupus-like disease and high interferon levels corresponding

to trisomy of the type I interferon cluster on chromosome 9p. Arthritis Rheum 54: 1573-1579.

31. Feng X, Petraglia AL, Chen M, Byskosh PV, Boos MD, et al. (2002) Low expression of interferon-stimulated genes in active multiple sclerosis is linked to subnormal phosphorylation of STAT1. J Neuroimmunol 129: 205-215.
32. Gresser I (1990) Biologic effects of interferons. J Invest Dermatol 95: 66S-71S.
33. Dupont SA, Goelz S, Goyal J, Green M (2002) Mechanisms for regulation of cellular responsiveness to human IFN-beta1a. J Interferon Cytokine Res 22: 491-501.
34. Lehtonen A, Matikainen S, Julkunen I (1997) Interferons up-regulate STAT1, STAT2, and IRF family transcription factor gene expression in human peripheral blood mononuclear cells and macrophages. J Immunol 159: 794-803.
35. Miller DH, Barkhof F, Frank JA, Parker GJ, Thompson AJ (2002) Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain 125: 1676-1695.
36. Rudick RA, Lee JC, Simon J, Fisher E (2006) Significance of T2 lesions in multiple sclerosis: A 13-year longitudinal study. Ann Neurol 60: 236-242.

Table S1. Gene details

Symbol	Name	Accession number GenBank
DHX58	DEXH (Asp-Glu-X-His) box polypeptide 58	NM_024119
EPSTI1	Epithelial stromal interaction 1 (breast)	NM_001002264
IFI27	Interferon, alpha-inducible protein 27	NM_005532
IFI44L	Interferon-induced protein 44-like	NM_006820
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	NM_001548
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	NM_001547
IMAGE:1926927		AI347124
IMAGE:2562181	2'-5'-oligoadenylate synthetase 2, 69/71kDa	NM_002535
IMAGE:504372		AA142842
IMAGE:545138		5'EST AA075776; 3'EST AA075725
IRF7	Interferon regulatory factor 7	NM_004031
ISG15	ISG15 ubiquitin-like modifier	NM_005101
LGALS3BP	Lectin, galactoside-binding, soluble, 3 binding protein	NM_005567
LY6E	Lymphocyte antigen 6 complex, locus E	NM_002346
MT1E	Metallothionein 1E	NM_175617
MT2A	Metallothionein 2A	NM_005953
MX1	Myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse)	NM_002462
OAS1	2',5'-oligoadenylate synthetase 1, 40/46kDa	NM_016816
OAS2	2'-5'-oligoadenylate synthetase 2, 69/71kDa	NM_002535
OAS3	2'-5'-oligoadenylate synthetase 3, 100kDa	NM_006187
PARP12	Poly (ADP-ribose) polymerase family, member 12	NM_022750
RSAD2	Radical S-adenosyl methionine domain containing 2	NM_080657
RTP4	Receptor (chemosensory) transporter protein 4	NM_022147
SAMD9L	Sterile alpha motif domain containing 9-like	NM_152703
SERPING1	Serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	NM_000062
TOR1B	Torsin family 1, member B (torsin B)	NM_014506
TRIM22	tripartite motif-containing 22	NM_006074
TRIM69	Tripartite motif-containing 69	NM_182985

Chapter 5. Summary and Discussion

5

Parts of this chapter are adapted from:
A genomic view of subtypes in rheumatoid arthritis: Towards personalized medicine van Baarsen E.G.M., van der Pouw-Kraan C.T.M, Verweij C.L.
Future Rheumatology June 2006 Vol. 1, No. 3, Pages 311-322

Transcription profiling of rheumatic diseases
van Baarsen E.G.M. , Bos C.L., Tineke van der Pouw-Kraan C.T.M, Verweij C.L.
Arthritis Res Ther. 2009 Jan 30;11(1):207

Pharmacogenomics of IFNβ in MS: Towards personalized medicine approach
Vosslamber S, van Baarsen E.G.M.and Verweij C.L.
Pharmacogenomics. 2009 Jan;10(1):97-108

The development of high-throughput techniques such as microarray analysis holds great promise for the unraveling of complex multifactorial diseases. After a decade of technical and analytical improvements, the technology and algorithms for data analysis have been shown to be robust and reproducible across properly designed and controlled experiments, and across different research groups. Encouraging results have been generated using microarray technology for the identification of predictors for disease outcome and metastasis and underlying pathways in breast cancer and lymphoma (1;2).
The results described in this thesis are focused on the autoimmune diseases RA and MS and highlight the importance of molecular redefinition of these diseases based on differential gene expression profiles, which provides a framework for further detailed investigation.

5.1 Genomics for early diagnosis of disease

5.1.1 Rheumatoid Arthritis (RA)

RA is one of the most common chronic inflammatory diseases for which no cure is available. Because of the destructive nature of the disease it would be highly desirable to start treatment as early as possible. Recognition of the preclinical phase allows a timely start of treatment with the ultimate goal of primary prevention. Several reports have described the appearance of autoantibodies against citrullinated proteins and RF years before the clinical symptoms of arthritis (3;4). In addition, a recent study showed that arthritis development in prospectively followed arthralgia patients is dependent on ACPA status and that autoantibodies can be present years without causing disease (5). Thus, autoantibodies may represent changes in the immune system that induce early pathogenesis. However, since not all autoantibody positive individuals are likely to develop arthritis we hypothesized that additional factors/processes are required or some individuals may have a protective immune response profile. To test this hypothesis we analyzed the peripheral blood gene expression profiles of autoantibody positive arthralgia patients without arthritis. In chapter 2 we identified four different molecular subgroups of autoantibody positive arthralgia patients associated with future development of arthritis. Interestingly, one subgroup of patients was associated with protection against arthritis development and characterized by an increased expression of genes involved in B-cell activity. The other subgroups were associated with development of arthritis and displayed increased expression levels of genes involved in IFN-mediated immunity, hematopoiesis and cytokine/chemokine mediated signaling. These results provide a framework for the discovery of biomarkers to predict the development of arthritis in autoantibody positive individuals at risk for RA.
The increased IFN response activity might have its origin in a disturbed response to an environmental trigger and/or an abnormality in genes involved in type I IFN

biology. Subsequently, the increased expression of genes involved in IFN-mediated immunity may affect the overall reactivity of the immune system and thus can predispose the individual to autoimmunity. This altered immunoreactivity is directed to particular antigens, i.e. citrullinated antigens, which affect B- and T-cell recognition of epitopes and autoantibody production. Genes involved in cytokine/chemokine mediated signalling may act on the progression of autoimmunity by modulating the immune response and altering cell trafficking to target tissues. Accordingly, Hueber and colleagues reported an increase in the levels of monocyte chemoattractant protein-1 (MCP-1) in ACPA positive patients before disease onset (6). Knowing that immune cell trafficking is a crucial process in the initiation of disease, it is tempting to speculate that the increased expression of B cell genes in individuals who did not develop arthritis is a reflection of a blockade in the migration of B cells to the target tissue.

On the basis of our data, we propose that the activity of individuals' genes affect susceptibility to RA at three levels. First, some genes affect the overall reactivity of the immune system and thus can predispose the individual to autoimmunity. Second, this altered immunoreactivity is directed to particular antigens and third, still other genes act on the progression of autoimmunity to target tissues to modulate immune attack. Our results imply that, among others IFN-mediated immunity and cell trafficking specify the processes relevant to progression to arthritis besides autoantibody positivity.

Additional specific processes that contribute to progression towards RA might involve epitope spreading. Mouse models of autoimmune arthritis and demyelination have shown the occurrence of epitope spreading to citrullinated antigens (7). In line with these results, a recent study showed a limited ACPA isotype usage in healthy relatives compared to RA patients. The authors speculate that epitope spreading and expansion of the isotype repertoire might be necessary for development of RA. It is tempting to speculate that the pathogenicity of the autoantibodies is related to the immune profile of the autoantibody positive individual. It will be interesting to determine if there is a relation between the different gene expression profiles in autoantibody positive arthralgia patients that we described and the fine specificity and isotype usage of their autoantibodies.

Perhaps our results in autoantibody positive arthralgia patients can be extrapolated to other diseases wherein autoantibodies are involved. In that case the production of autoantibodies in individuals who have a skewed immune response profile may lead to epitope spreading, immune activation, cell trafficking and development of disease. Therefore the combined analysis of different autoantibodies together with gene expression profiles may be predictive for future development of different autoimmune diseases. Future studies are needed to test this hypothesis.

5.1.2 Multiple Sclerosis (MS)

The prevalence of MS is much lower than RA and diagnosis of disease is generally at a younger age, which makes large studies using blood donors like those performed in RA more difficult. Unfortunately, preclinical markers like RF and ACPA have not yet been identified in MS patients. However, a recent nested case-control study of US military personnel (63% male) who have serum samples stored in the Department of Defense Serum Repository showed that the presence of anti-MOG (myelin oligodendrocyte glycoprotein) IgM-/IgG+ antibodies measured by ELISA was associated with a moderately increased risk of developing MS (8). Interestingly, since this association was lost after adjustment for antibody titers to Epstein-Barr nuclear antigens (EBNA), the association may in part reflect cross-reactivity between MOG and EBNA.

5.2 Genomics for patient subclassification

MS and RA are both complex chronic inflammatory diseases with unknown etiology and a multifactorial background. Gene expression profiling of peripheral blood cells of RA and MS patients revealed a remarkable variation in gene expression levels leading to subclassification of patients depending on their molecular signature.

5.2.1 Heterogenomics in RA

A large-scale gene expression profiling study of synovial tissues from patients with erosive RA revealed considerable heterogeneity among different patients (9;10). It has been demonstrated that different arthroscopic biopsies taken from one joint result in gene expression signatures that are more similar within the joint of one patient than between patients (11). A systematic characterization of the differentially expressed genes highlighted the existence of at least two molecular distinct forms of RA tissues. One group revealed abundant expression of clusters of genes indicative of an ongoing inflammation and involvement of the adaptive immune response. This subgroup is referred to as the RA high inflammation group. Further analyses of the genes involved in the high inflammation tissues provide evidence for a prominent role of genes indicative for an activated STAT-1 pathway. These findings were confirmed at the protein level (12;13)

Tsubaki and colleagues demonstrated that tissue heterogeneity within RA can already be observed in the early phase of RA (14). In this study, gene expression profiles were analyzed from synovial lining tissues from 12 patients with early RA (duration of less than 1 year after diagnosis) and 4 with long-standing RA (duration of more than 3 year after diagnosis). As seen in the previous study using biopsies from long-standing RA patients, the early RA patients could be divided in at least two different groups based on their gene expression profiles.

In chapter 3 we confirmed and further extended the molecular heterogeneity at

tissue level using synovial tissue biopsies obtained during arthroscopic surgery. In addition, the subclassification of synovial tissues based on gene expression profiles matched the subclassification based on immunohistochemical data using protein markers for infiltrating cells and inflammatory cytokines. The high inflammatory tissues were characterized by a high expression of genes involved in many inflammation-mediated processes such as immunity and defense and T and B cell mediated immunity. In contrast, the low inflammatory tissues displayed increased expression of genes involved in amongst others developmental processes and neurogenesis. We demonstrated the clinical relevance of tissue subtyping. RA patients with a high inflammatory subtype had a higher disease activity score (DAS) together with a shorter disease duration compared to patients with a low inflammatory tissue type. Since data on normal healthy synovial tissue was not available we were not able to investigate how these biological activities relate to "normal" levels.

The existence of molecular heterogeneity in rheumatoid synovial tissues fits a model proposed by Firestein and Zvaifler (15), who suggested two independent phases that drive destruction of bone and cartilage; an antigen-driven immune cell mediated process and a subsequent antigen-independent stromal cell driven process. This model is extended by Kirwan who proposed the existence of two pathogenic mechanisms in which synovial inflammation will lead to joint pain, joint swelling and cartilage thinning, whereas synovial hyperplasia will lead to joint swelling and erosions (16). Accordingly, we showed that the high inflammatory tissue type is accompanied with a higher tender joint count. Since our patient groups also differ in disease duration, the observed tissue heterogeneity may also reflect two different phases of joint destructions; an acute antigen-dependent inflammatory cell type phase and a chronic stromal cell type driven phase. The transition from immune-mediated inflammation to a more autonomous, immune-independent process may occur several years after onset of clinical signs and symptoms. Longitudinal studies are needed to test if these pathogenic mechanisms are running in parallel or in time.

▪ 5.2.1.1 Heterogenomics of peripheral blood cells from RA patients

While gene expression analysis of synovial tissue samples of affected organs offers insights into the genes that are instrumental in patient stratification and primarily involved in disease activity and pathogenesis, this approach is not feasible in studies of large cohorts of patients. Due to the systemic nature of RA and the communication between the systemic and organ specific compartments, we and others have studied whole blood and/or peripheral blood mononuclear cells to obtain disease-related gene expression profiles. The peripheral blood is especially suitable to analyze gene expression profiles that can be used as biomarker to select patients for improved diagnosis as well as for patient centric therapy approach.

In chapter 3 we performed gene expression profiling analysis on whole blood cells,

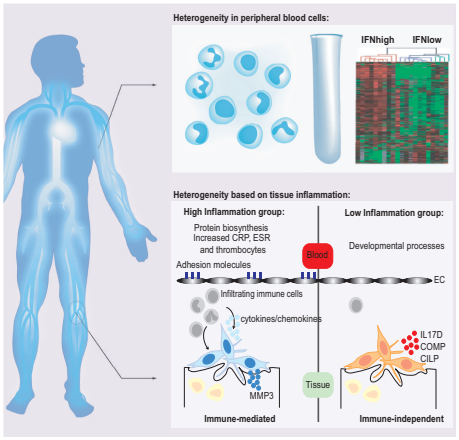
which revealed clear and significant differences between RA patients (n=35) and healthy individuals (n=15). The microarray data confirmed previous observations of increased expression of the calcium-binding protein S100A8 and S100A12. Application of gene set analysis algorithms to identify gene sets that reflect a distinct pathway or biological process revealed an increased expression of immune defense genes, including type I IFN-response genes, suggesting that this pathway is activated systemically. Most interestingly, the increased expression of immune defense genes was characteristic of not all but approximately half of the patients (17). In addition, a combined cluster analysis of RA and control samples together with samples from a viral infection model revealed that the gene expression profile of a subgroup of RA patients was reminiscent to that of poxvirus-infected macaques (18). Based on these findings we conclude that activation of the immune response, with among the gene sets a type I IFN signature, defines a subgroup of RA patients. Similar findings were reported by Olsen and colleagues who studied differentially expressed genes between early (disease duration less than 2 years) and established RA (with an average disease duration of 10 years) (19). Genes involved in immune/ inflammatory processes and genes related to cell proliferation and neoplasia were expressed at lower levels in early arthritis. About a quarter of the early arthritis genes overlapped with an influenza-induced gene set. This finding led the authors to suggest that the early arthritis signature may partly reflect the response to an unknown infectious agent. Together these results are supportive for differential expression of IFN response gene activity in RA, which could be the consequence of either an exogenous or endogenous stimulatory agent.

5.2.1.2 Linking tissue heterogeneity to peripheral blood profiles

In chapter 3.3 we explored the presence of blood based markers associated with tissue inflammation. Whereas the synovial tissues clearly demonstrated the presence of a high and low inflammatory tissue type, this difference was not found to be linked to the paired peripheral blood gene expression level. These results imply that the reflection of tissue inflammation in the peripheral blood cell compartment is not that obvious. We confirmed other studies by showing that patients with a higher DAS and higher level of tissue inflammation display increased levels of CRP combined with a higher ESR and more thrombocytes. The lack of molecular peripheral blood cell activation markers measured at transcriptional level in combination with increased CRP, ESR and thrombocytes in the blood maybe partially explained by the fact that CRP is an acute phase protein produced in the liver, thrombocytes have negligible amounts of RNA and the finding that the frequency of white blood cells were not different between the two groups. In addition, we should keep in mind that we analyzed the gene expression profiles of only one affected joint which may explain the difficulties in finding correlating blood markers since inflammatory processes may be different between joints.
Using a pathway level analysis approach we found that the peripheral blood cells

of patients with a high inflammatory tissue type displayed increased expression of genes involved in protein biosynthesis. Conversely, peripheral blood cells of patients with a low inflammatory tissue type showed an enriched expression of genes involved in developmental processes. Further studies are needed to elucidate the role of these biological processes in tissue inflammation. Figure 1 summarizes the heterogenomics of RA.

Figure 1. Discovery of RA subtypes in peripheral blood cells and synovial tissue



Schematic overview of the discovery of RA subtypes in peripheral blood cells and synovial tissue. Hierarchical clustering of differentially expressed genes in PB cells and synovial tissues from RA patients highlighted the existence of at least two molecular distinct forms of RA. It is hypothesized that two different pathogenic mechanisms are playing a role in RA. The first is an immune cell mediated destruction of cartilage via activation of matrix metalloproteases such as MMP3 whereas the second is an immune independent mechanism possible mediated by tissue remodeling and growth factors. Patients with a high inflammatory tissue type display increased CRP, ESR and thrombocyte counts. In addition, a subgroup of patients has increased expression levels of type I IFN response genes. (Adapted and changed from Future Rheumatology June 2006 Vol. 1, No. 3, Pages 311-322, with permission of Future Medicine Ltd)

5.2.2 Heterogenomics in MS

Since it is highly difficult to obtain from large study groups brain tissue during the course of disease it is not easy to elucidate the role of tissue heterogeneity in the pathogenesis of disease. Because of the autoimmune nature of the disease and the migration of immune cells to the brain, we hypothesized that the disease processes are reflected in the peripheral blood compartment which we are able to study during disease.

In chapter 3.4 we performed gene expression profiling in peripheral blood (PB) cells of MS patients which revealed an elevated expression of a spectrum of genes involved in immune defense e.g. TLR signaling when compared to healthy controls. Most strikingly, our results demonstrated evidence for a remarkable heterogeneity between the RRMS patients. This was primarily based on differential expression of type IFN response genes. Moreover, the expression signatures of half of the MS patients were similar to those of virus-infected macaques (20). Collectively, these findings provide a basis for patient stratification in RRMS and may support the hypothesis that viral antigens play an important role in the pathogenesis MS. However, since endogenous antigens such as DNA or RNA containing immune complexes can also activate certain TLR signaling routes leading to increased

transcription of type I IFN induced genes (21), further studies are needed to elucidate the origin of the type I IFN induced gene expression signature. Similar results have been generated by Yamaguchi and colleagues who also showed a dysregulated expression of IFNβ response genes in treatment naïve RRMS patients (22).

5.3 Common denominators in autoimmunity; what's the role of IFN?

Upregulation of IFN-induced genes has now been observed in PB cells and/or target tissue of (a subset of) patients with different autoimmune diseases e.g. RA, SLE, SSc, SS, MS, and type 1 diabetes. These findings suggest that an activated IFN gene expression program may be a common denominator in autoimmune diseases in general.

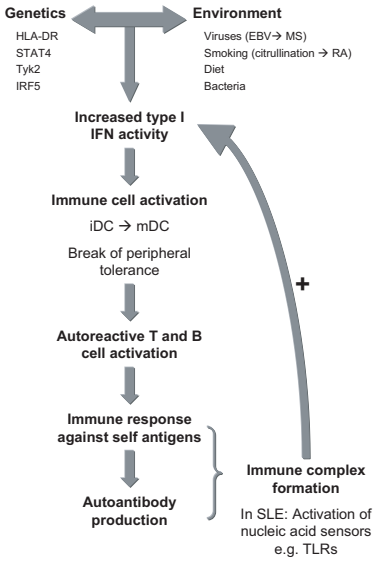


Figure 2. The possible role of IFN in autoimmunity
Both genetic and environmental factors may lead to an increased expression of type I IFN-response genes. The increased activity of type I IFN-response genes may lead to accelerated DC activation and a break of peripheral tolerance resulting in activation of autoreactive T and B cells. The subsequent immune response against self-antigens will lead to autoantibody production. In SLE the autoantibody production may lead to immune complex formation which can activate immune cells to produce more type I IFN eventuating in a vicious circle of immune activation and IFN production.

The initial activation of the type I IFN gene activity most likely depends on both the genetic background as well as encountered environmental (endogenous and/or exogenous) factors of the individual. Individuals harboring a certain genetic profile of e.g. risk alleles for genes involved in IFN biology may than generate increased levels of IFN-response genes. It is speculated that the IFN-response program could be associated with the activation of immature DCs, which regulate deletion of autoreactive lymphocytes. Subsequently, IFN-matured DCs may activate autoreactive T cells leading to autoreactive B cell development and a first level of autoimmunity (23). Loss of tolerance for self antigens may then lead to autoantibody production.

In the case of SLE, autoantigen/autoantibody complexes may trigger pathogen recognition receptors (such as TLRs) that induce IFNα production and thereby perpetuates the IFN response program. Consequently, a vicious circle is established of increased type I IFN-response activity and autoreactivity against self antigens resulting in autoimmune disease (Figure 2). Breaking

this circle resulting in downregulation of the type I IFN-response gene activity could be a therapeutic approach for SLE. Indeed, recent results from a phase I clinical trial showed reduction in clinical disease activity and downregulation of type I IFN-response genes in SLE patients treated with an anti-IFNα mAb (24). Conversely, type I IFN treatment is beneficial in MS patients and in animal models of arthritis.

Besides a role for the IFN response program as common denominator in autoimmune diseases other gene expression profiles have been reported that are shared by autoimmune diseases. In particular Maas and colleagues (25) studied the overlap of gene expression profiles between different diseases. They identified 95 genes that were increased and 117 genes that were decreased in PBMC of all patients with RA, SLE, type I diabetes and MS. These genes were involved in e.g. inflammation, signaling, apoptosis, ubiquitin/proteasome function and cell cycle. Hierarchical cluster analysis on the basis of gene signatures in PBMC revealed that RA and SLE patients were intermixed with one another. Moreover, they reported that from the genes that were differentially expressed between PBMCs from patients and unrelated unaffected individuals, the gene expression profile of 127 genes was shared between patients with autoimmune diseases and unaffected first-degree relatives. This commonality between affected and unaffected first degree relatives suggests a genetic basis for these shared gene expression profiles. Accordingly, they showed that these genes are clustered in chromosomal domains supporting the hypothesis that there is some genetic logic to this commonality (26).

5.4 Pharmacogenomics

Given the destructive nature of autoimmune diseases it would be highly desirable to predict in an early stage the most beneficial treatment for the patients at risk. If we rely solely on clinical or radiographic/MRI manifestations we will probably be responding too late in order to maximize protection. Ideally, it would be desirable to make predictions on the success before the start of therapy. Ultimately, this may lead to a personalized form of medicine, whereby a specific therapy will be applied that is best suited for an individual patient.
We applied genomics for studying the pharmacodynamics of IFNβ and anti-TNF therapy and for predicting the response to therapy. The term *pharmacogenomics* emerged in the late 1990s and is associated with the application of genomics in drug development. *Pharmacogenomics* is defined as: "The investigation of variations of DNA and RNA characteristics as related to drug response". In this thesis we focused on transcriptomics which investigates RNA characteristics.

5.4.1 Pharmacogenomics of anti-TNF treatment in RA patients

TNF antagonists are approved worldwide for the treatment of various rheumatic

diseases. Clinical experience indicates that there are 'responders' as well as 'non responders', however clear criteria for such classification are still lacking. For RA, treatment is only effective for approximately two-third of the patients (27) which has attracted interest in the pharmacology and mechanism of action.
Until now a few pharmacogenomics studies have been published to gain insight in the pharmacodynamics and to identify genes predictive of responsiveness to TNF blockers. Thus far, studies aimed at identifying biomarkers before the treatment to predict the response to anti-TNF treatment in RA have not revealed consistent results. The discrepancies are partly explained because of differences in study design, experimental platform and patients included. In addition, universal criteria for defining the clinical response to treatment are currently lacking. Therefore, additional studies using large cohorts of patients and more stringent response criteria are necessary.
In chapter 4.1 we describe the pharmacogenomics of Infliximab treatment in RA patients (n=15) before and 1 month after the start of infliximab. These results revealed a similar change in the expression of a pharmacogenomic response gene set in the peripheral blood compartment of all patients treated irrespective of clinical response. This study indicates that all RA patients exhibit an active TNF response program contributing to disease pathogenesis (28). However, since not every RA patient shows a good clinical response to anti-TNF treatment we hypothesize that in these non-responder patients additional biological programs besides TNF are involved in disease pathogenesis. The relative contribution of these TNF-independent pathways may determine clinical response to treatment.

Although TNF antagonists have been shown to be effective for the treatment of several inflammatory diseases including RA, several clinical complications have been reported. Patients treated with TNF antagonists may develop lupus-like symptoms and increased dsDNA autoantibody levels. It was hypothesized that these complications were the result of anti-TNF induced increase in type I IFN

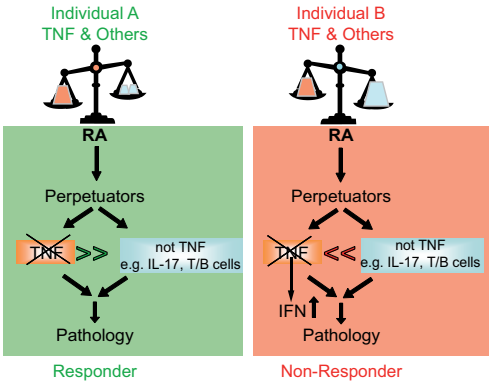


Figure 3. Pharmacogenomics of anti-TNF treatment in RA
By analyzing the change in gene expression profiles upon antagonizing the function of TNF we demonstrated that all RA patients exhibit an active TNF response program contributing to disease pathogenesis. The results described in chapter 4 favour a model for the parallel presence of TNF-dependent and TNF-independent pathways (named "others") in the individual RA patient. Clinical response to TNF blockade might be determined by the relative contribution of the TNF-independent pathways to disease. Moreover, poor response to treatment is associated with TNF blockade induced expression of type I IFN response genes.

activity since SLE is characterized by increased expression levels of IFN-response genes in PB and because SOJIA patients treated with TNF antagonist have increased expression levels of IFN-response genes (29). However, as described in chapter 4.2, gene expression studies in whole blood of RA patients before and 1, 2, 3 and 4 months after the start of TNF-blockade (infliximab) revealed a variable effect on the expression of IFN-response genes upon treatment. Therefore, the positive effect of TNF blockade on IFN is not consistently observed in RA. The anti-TNF induced increase in type I IFN gene activity was not associated with the extent of IFN-response gene activity at baseline. Interestingly, the anti-TNF induced increase in type I IFN-response gene activity was associated with poor clinical response to treatment as assessed by EULAR criteria. Although the anti-TNF induced increase in IFN-response gene activity might be an epiphenomenon related to the effect of TNF blockade, it is tempting to speculate on a role of increased IFN bioactivity in the deteriorating clinical effects (as discussed in chapter 5.3). Figure 3 summarizes our main findings with respect to pharmacogenomics of TNF blockade in RA.

▪ 5.4.2 Pharmacogenomics of IFNβ treatment in MS

Interferons (IFN) were the first agents to show clinical efficacy in treatment of MS, and prolonged treatment is still the best available therapy. Because MS is a chronic disease with an unpredictable clinical course it remains difficult to assess clinical responder status at an individual patient level. Despite availability of objective methods for determining disease activity such as the measurement of MRI parameters, e.g. CNS atrophy measures or T1 gadolinium enhancing or the appearance of new T2 lesions (30-32) it is still extremely difficult to precisely define the state of responsiveness after a short period of treatment. This emphasizes the importance of finding pharmacological response markers that could aid in the assignment of clinical responders and non-responders of IFNß therapy.

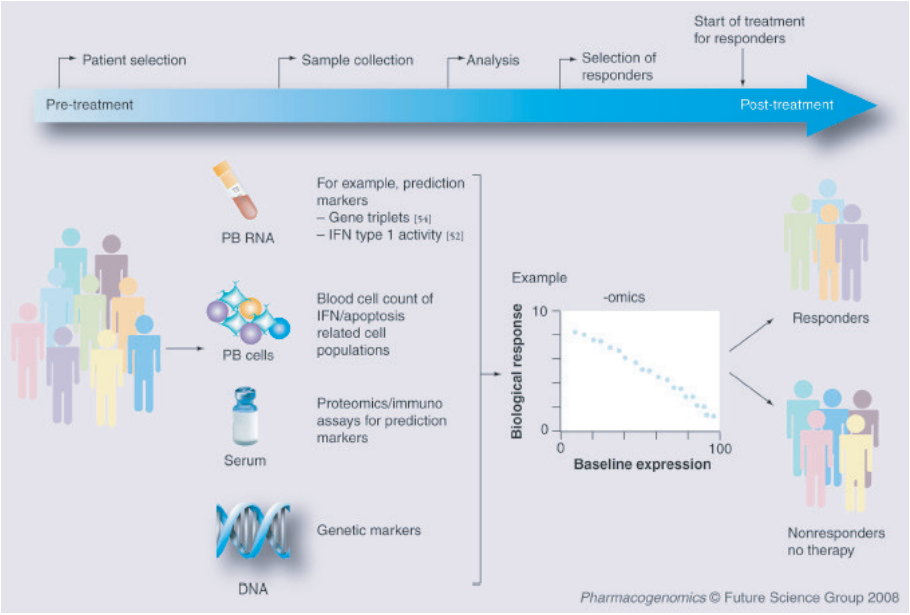
In chapter 4.3 we described the change in gene expression profiles in PB after IFNβ therapy, or in other words: the pharmacogenomics of IFNβ treatment in PB. The heterogeneous expression of type I IFN induced genes in untreated MS patients was confirmed and we investigated how this signature was related to treatment response. As anticipated (22), type I IFN response genes were increased upon treatment. Interestingly, it was observed that the extent of the pharmacological response correlates negatively with the baseline expression of a specific set of IFN response genes. Furthermore, these *ex vivo* results could be predicted *in vitro* using IFNβ stimulated PB cells isolated prior to treatment. Overall, we concluded that the inter-individual variation in response to IFNβ therapy is an intrinsic property of the PB cell compartment and unrelated to treatment regimens and interfering serum proteins such as NAbs.

Because MS is a chronic disease with an unpredictable disease course it is highly difficult to determine the clinical response to treatment at the individual patient level. Due to the small sample size of our study we lacked power to address the

question of an association with the clinical response to treatment. Nonetheless, these data imply that the type I IFN induced gene expression signature prior to treatment determines the pharmacological difference between patients and could serve a role as biomarker for clinical response to IFNβ. Currently, experimental studies are carried out with larger study groups and well defined patient samples to investigate the clinical relevance of the observed results.

The inter-individual differences in IFNß responsiveness may be the result of genetic variation in the IFNß biology. Other mechanisms that could account for differential responsiveness to IFNß include variation in activity of inhibitory transcription factors. Evidence exists that crosstalk with other cytokine-activated pathways, could cause tachyphylaxis to type I IFNs. Although type I IFNs have an essential function in mediating innate immune responses against viruses, they may already be produced at very low levels in the absence of viral infections (33) in serum of a subset of MS patients. Since e.g. IFNα is known to desensitize further responses to IFNs (62), the presence of low endogenous IFNs could block IFNß-induced signals (34;35). Future studies aimed at unraveling these possibilities should shed new light on the mechanism of variation in IFN response in relation to the endogenous IFN signature. Figure 4 describes the future activities towards a personalized medicine approach for IFNß in RRMS.

Figure 4. Steps towards a personalized medicine approach for IFNβ therapy in RRMS



(Reproduced from Pharmacogenomics 2009 Jan;10(1):97-108, with permission of Future Medicine Ltd)

5.5 Conclusion

Taken together, the work described in this thesis has provided novel information about molecular heterogeneity in relation to disease development, disease subclassification, disease pathogenesis and response to treatment. Accordingly, evidence is accumulating that large-scale gene expression profiling will contribute considerably to our understanding of the molecular and biological basis of the well-recognized but as yet poorly defined heterogeneity of diseases such as MS and RA.

Clearly, these developments open the way for a "redefinition" of diseases like RA and MS. The molecular markers (biomarkers) that distinguish patients from one another can be used for early diagnosis, prognosis, determining drug efficacy and risk management. The transition of this explorative type of research to integrated clinical diagnostics will require extensive optimization of the technological procedures. Moreover, the assignment and implementation of useful and reliable classifiers requires rigorous standardization and several levels of validation.

Hence future treatment will be tailored to the molecular and biological features of an individual patient. This will lead to personalized medicine, i.e. the prescription of therapeutics that is best suited for an individual patient.

5.6 Reference List

(1) Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000; 403(6769):503-511.

(2) van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; %19;347(25):1999-2009.

(3) Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004; 50(2):380-386.

(4) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003; 48(10):2741-2749.

(5) Bos WH, Boers M, Wolbink G, Tijhuis G, de Vries N, Bruinsma IV et al. Arthritis development in prospectively followed arthralgia patients is dependent on anti-citrullinated protein antibody status. Arthritis Rheum 2008; 58(9):S290.

(6) Hueber W, Tomooka BH, Zhao X, Kidd BA, Drijfhout JW, Fries JF et al. Proteomic analysis of secreted proteins in early rheumatoid arthritis: anti-citrulline autoreactivity is associated with up regulation of proinflammatory cytokines. Ann Rheum Dis 2007; 66(6):712-719.

(7) Kidd BA, Ho PP, Sharpe O, Zhao X, Tomooka BH, Kanter JL et al. Epitope spreading to citrullinated antigens in mouse models of autoimmune arthritis and demyelination. **Arthritis Res Ther 2008; 10(5):R119.**

(8) Wang H, Munger KL, Reindl M, O'Reilly EJ, Levin LI, Berger T et al. **Myelin oligodendrocyte glycoprotein antibodies and** multiple sclerosis in healthy young adults. Neurology 2008; 71(15):1142-1146.

(9) van der Pouw Kraan TC, van Gaalen FA, Huizinga TW, Pieterman E, Breedveld FC, Verweij CL. Discovery of distinctive gene expression profiles in rheumatoid synovium using cDNA microarray technology: evidence for the existence of multiple pathways of tissue destruction and repair. Genes Immun 2003; 4(3):187-196.

(10) van der Pouw Kraan TC, van Gaalen FA, Kasperkovitz PV, Verbeet NL, Smeets TJ, Kraan MC et al. Rheumatoid arthritis is a heterogeneous disease: evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. Arthritis Rheum 2003; 48(8):2132-2145.

(11) Lindberg J, af KE, Ulfgren AK, Stark A, Andersson T, Nilsson P et al. Variability in synovial inflammation in rheumatoid arthritis investigated by microarray technology. Arthritis Res Ther 2006; 8(2):R47.

(12) Kasperkovitz PV, Verbeet NL, Smeets TJ, van Rietschoten JG, Kraan MC, van der Pouw Kraan TC et al. Activation of the STAT1 pathway in rheumatoid arthritis. Ann Rheum Dis 2004; 63(3):233-239.

(13) Lorenz P, Ruschpler P, Koczan D, Stiehl P, Thiesen HJ. From transcriptome to proteome: differentially expressed proteins identified in synovial tissue of patients suffering from rheumatoid arthritis and osteoarthritis by an initial screen with a panel of 791 antibodies. Proteomics 2003; 3(6):991-1002.

(14) Tsubaki T, Arita N, Kawakami T, Shiratsuchi T, Yamamoto H, Takubo N et al. Characterization of histopathology and gene-expression profiles of synovitis in early rheumatoid arthritis using targeted biopsy specimens. Arthritis Res Ther 2005; 7(4):R825-R836.

(15) Firestein GS, Zvaifler NJ. How important are T cells in chronic rheumatoid synovitis?: II. T cell-independent mechanisms from beginning to end. Arthritis Rheum 2002; 46(2):298-308.

(16) Kirwan JR. The synovium in rheumatoid arthritis: evidence for (at least) two pathologies. Arthritis Rheum 2004; 50(1):1-4.

(17) van der Pouw Kraan TC, Wijbrandts CA, van Baarsen LG, Voskuyl AE, Rustenburg F, Baggen JM et al. Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: assignment of a type I interferon signature in a subpopulation of patients. Ann Rheum Dis 2007; 66(8):1008-1014.

(18) van der Pouw Kraan TC, van Baarsen LG, Wijbrandts CA, Voskuyl AE, Rustenburg F, Baggen JM et al. Expression of a pathogen-response program in peripheral blood cells defines a subgroup of Rheumatoid Arthritis patients. Genes Immun 2007.

(19) Olsen N, Sokka T, Seehorn CL, Kraft B, Maas K, Moore J et al. A gene expression signature for recent onset rheumatoid arthritis in peripheral blood mononuclear cells. Ann Rheum Dis 2004; 63(11):1387-1392.

(20) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006; 7(6):522-531.

(21) Lovgren T, Eloranta ML, Bave U, Alm GV, Ronnblom L. Induction of interferon-alpha production in plasmacytoid dendritic cells by immune complexes containing nucleic acid released by necrotic or late apoptotic cells and lupus IgG. Arthritis Rheum 2004; 50(6):1861-1872.

(22) Yamaguchi KD, Ruderman DL, Croze E, Wagner TC, Velichko S, Reder AT et al. IFN-beta-regulated genes show abnormal expression in therapy-naive relapsing-remitting MS mononuclear cells: gene expression analysis employing all reported protein-protein interactions. J Neuroimmunol 2008; 195(1-2):116-120.

(23) Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity 2006; 25(3):383-392.

(24) Wallace DJ, Petri M, Olsen N, Kirou K, Dennis G, Yao Y et al. MEDI-545, an Anti-Interferon Alpha Monoclonal Antibody, Shows Evidence of Clinical Activity in Systemic Lupus Erythematosus. Arthritis and Rheumatism 2007, S526-S527. 2007. Ref Type: Abstract

(25) Maas K, Chen H, Shyr Y, Olsen NJ, Aune T. Shared gene expression profiles in individuals with autoimmune disease and unaffected first-degree relatives of individuals with autoimmune disease. Hum Mol Genet 2005; 14(10):1305-1314.

(26) Aune TM, Maas K, Parker J, Moore JH, Olsen NJ. Profiles of gene expression in human autoimmune disease. Cell Biochem Biophys 2004; 40(2):81-96.

(27) Maini R, St Clair EW, Breedveld F, Furst D, Kalden J, Weisman M et al. Infliximab (chimeric anti-tumour necrosis factor alpha monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomised phase III trial. Lancet 1999; 354(9194):1932-1939.

(28) van Baarsen EGM, Wijbrandts CA, Rustenburg F, van der Pouw Kraan TCTM, Dijkmans BAC, Tak PP et al. Pharmacogenomics

of Anti-TNF Treatment in Rheumatoid Arthritis Reveals an Active Baseline TNF Response Profile in All Patients. Arthritis and Rheumatism , *in press*. 2008 Ref Type: Abstract

(29) Palucka AK, Blanck JP, Bennett L, Pascual V, Banchereau J. Cross-regulation of TNF and IFN-alpha in autoimmune diseases. Proc Natl Acad Sci U S A 2005; 102(9):3372-3377.

(30) Rudick RA, Lee JC, Simon J, Ransohoff RM, Fisher E. Defining interferon beta response status in multiple sclerosis patients. Ann Neurol 2004; 56(4):548-555.

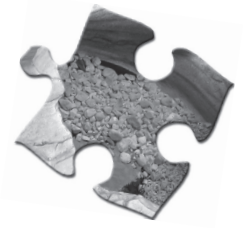
(31) Rudick RA, Lee JC, Simon J, Fisher E. Significance of T2 lesions in multiple sclerosis: A 13-year longitudinal study. **Ann** Neurol 2006; 60(2):236-242.

(32) Miller DH, Barkhof F, Frank JA, Parker GJ, Thompson AJ. **Measurement of atrophy in multiple sclerosis: pathological basis,** methodological aspects and clinical relevance. Brain 2002; 125(Pt 8):1676-1695.

(33) Dupont SA, Goelz S, Goyal J, Green M. Mechanisms for regulation of cellular responsiveness to human IFN-beta1a. J Interferon Cytokine Res 2002; 22(4):491-501.

(34) Gresser I. Biologic effects of interferons. J Invest Dermatol 1990; 95(6 Suppl):66S-71S.

(35) Lehtonen A, Matikainen S, Julkunen I. Interferons up-regulate STAT1, STAT2, and IRF family transcription factor gene expression in human peripheral blood mononuclear cells and macrophages. J Immunol 1997; 159(2):794-803.



Chapter 6

Dutch summary - Nederlandse samenvatting voor niet ingewijden
Acknowledgements - Dankwoord
Curriculum Vitae
List of publications

6

Nederlandse samenvatting voor niet ingewijden

In een goed functionerend lichaam worden immuun cellen, die lichaamseigen antigenen herkennen, al in een vroeg stadium geëlimineerd. Indien dit niet gebeurd kan er een ongewenste immune reactie gestart worden tegen ons eigen lichaam, wat kan leiden tot de ontwikkeling van een auto-immuun ziekte. Auto-immune ziektes worden gekenmerkt door lokale weefselschade, veroorzaakt door de continue infiltratie en activering van ontstekingscellen. Deze geactiveerde ontstekingscellen produceren ontstekingsstimulerende en/of ontstekingsremmende eiwitten (zgn. cytokines en chemokines) die effect hebben op de weefselschade en de rekrutering van nieuwe ontstekingscellen. De balans tussen de verschillende immuun modulerende cytokines is van groot belang en kan bepalend zijn voor de ontwikkeling van een beschermende of juist een overactieve immune reactie. Cytokines die een belangrijke rol hebben bij auto-immuun ziektes zijn tumor necrose factor (TNF) en type I interferon (IFN). Reumatoïde artritis (RA) en multiple sclerose (MS) zijn beide auto-immuun ziektes waarbij verschillende cytokines een rol spelen. Zo is het remmen van TNF een goede therapie voor patiënten met RA, terwijl dezelfde therapie voor patiënten met MS de ziekte verergert. Voor patiënten met MS geeft een behandeling met type I IFN juist een gunstig effect.

Reumatoïde artritis (RA)

RA is een chronische gewrichtsaandoening die naar schatting 1% van de wereldbevolking treft. Patiënten met RA hebben ontstekingsreuma dat zich richt op het beschadigen van bot en kraakbeen, wat uiteindelijk leid tot pijnlijke en opgezwollen gewrichten. De oorzaak van de ziekte is onbekend, maar zowel erfelijke als omgevingsfactoren lijken een belangrijke rol te spelen in het ontstaan van de ziekte. Tegenwoordig zijn er wereldwijd gebruikte criteria voor het stellen van de diagnose RA, maar genezen is tot op heden niet mogelijk. Omdat niet iedere RA patiënt dezelfde symptomen en klachten vertoont, is het ziekteverloop onvoorspelbaar en zeer afwisselend tussen patiënten en kunnen verschillende behandelingstrategieën nodig zijn. Voorts is het belangrijk zo snel mogelijk na diagnose te starten met behandelen aangezien er anders onomkeerbare gewrichtsschade kan ontstaan. Onderzoek gedurende de laatste decennia heeft geleid tot de ontwikkeling van zogenaamde biologicals die de werking van lichaamseigen stoffen van het immuunsysteem nabootsen of beïnvloeden. Een veel toegepaste biological in RA is anti-TNF dat het cytokine TNF wegvangt en zo de ontsteking snel en krachtig onderdrukt. De beschikbare medicijnen kunnen het ziekteproces remmen, maar werken helaas niet bij alle patiënten met RA. Het vinden van nieuwe medicijnen en het vroeg diagnosticeren van patiënten is daarom ontzettend belangrijk. Bovendien zou het van onschatbare waarde zijn om vóór de start van een behandeling al te kunnen voorspellen of de patiënt zal gaan reageren op een bepaalde therapie. Dit spaart immers tijd, geld en eventuele nare bijwerkingen voor de patiënt.

Multiple sclerose (MS)

MS is een ziekte van het centrale zenuwstelsel (hersenen en ruggenmerg) en wordt gedacht van oorsprong een auto-immuun ziekte te zijn. De ziekte begint op een relatief jonge leeftijd (tussen de 25-45 jaar) en is op dit moment niet te genezen. Het is nog onduidelijk waarom bepaalde mensen MS ontwikkelen. Wel wordt dankzij onderzoek steeds meer duidelijk dat het ontstaan van deze ziekte te maken heeft met het verstoord werken van de regulerende functie van de afweer (immuniteit). Myeline is een isolatie laag van axonen (zenuwbanen) en vergroot de geleidingssnelheid van zenuwimpulsen. In MS patiënten is deze myeline laag beschadigd (laesie), waardoor de snelheid van geleiding over de axonen wordt verminderd. Dit kan leiden tot neurologische symptomen (bv. coördinatie stoornissen, slecht zien, spierzwakte en vermoeidheid) die afhangen van de plaats van de laesie. De symptomen van MS zijn vaag, kunnen verschillend zijn tussen patiënten en kunnen ook bij andere ziektes optreden waardoor het stellen van de diagnose complex is. Bovendien is het verloop van de ziekte afwisselend en patiënten reageren verschillend op behandeling met IFNβ, de meest voorgeschreven biological voor MS patiënten. MS wordt hierdoor net als RA beschouwd als een complexe heterogene ziekte.

Moleculaire markers om verschillende (klinische) vormen van MS te classificeren zouden een enorme bijdrage kunnen leveren voor diagnose en behandeling. Bovendien zouden deze markers kunnen helpen bij het definiëren van homogene patiëntengroepen voor het opzetten van klinische studies naar nieuwe behandelingsmethoden. Deze homogene groepen zal de kans op het bereiken van doeltreffendheid van nieuwe behandelingswijzen vergroten.

Het promotieonderzoek

MS en RA zijn beide heterogene ziektes, wat onder andere blijkt uit de verscheidenheid in klinische presentatie van de ziekte en de grote verschillen in respons op therapie. De oorzaak van beide ziektes is onbekend en ons inzicht in het ziekteproces is beperkt. Goede criteria om verschillende vormen van MS of RA te onderscheiden zijn momenteel niet voorhanden. Het kunnen definiëren van homogene patiëntengroepen is van essentieel belang voor het uitvoeren van klinische studies en het verhogen van de kans op succes bij het testen van nieuwe medicijnen. Daarom richt dit proefschrift zich ten eerste op het hergroeperen van patiënten met een chronisch ontstekingsziekte in meer homogene patiënten groepen.

Het ontstaan en beloop van ziektes gaat gepaard met aanzienlijke veranderingen in de activiteit van onze genen. De hoeveelheid genproducten (genexpressie) en de complexe interactie tussen deze genproducten worden bepalend geacht voor de ziekteactiviteit. Recent heeft een revolutionaire ontwikkeling in de technologie van het meten van genactiviteit plaatsgevonden die het mogelijk maakt de genexpressie van tienduizenden genen tegelijkertijd te bepalen. Dit wordt

gedaan met behulp van een van de nieuwste technologische ontwikkeling op het moleculaire gebied, de zogenaamde DNA microchips. Deze chip meet in één experiment de activiteit van ongeveer 30,000 genen, de boodschappers van ons DNA. Deze informatie kan per patiënt worden omgezet in een soort streepjescode waarna we de meest op elkaar lijkende codes bij elkaar zoeken en in één groep plaatsen. Per patiëntengroep kunnen we vervolgens de activiteit en de onderlinge connectie van de genen bestuderen om zo het onderliggende ziekteproces te ontrafelen en nieuwe targets voor therapieën te identificeren. Dit soort studies wordt ook wel genomics (studie van het genoom) onderzoek genoemd.

Uiteindelijk willen we middels genomicsonderzoek komen tot eenduidige diagnostische en prognostische markers die het mogelijk maken om binnen de heterogene ziektes te komen tot het samenstellen van homogene patiëntenpopulaties voor klinische studies en behandeling. Voorts biedt het gedetailleerd inzicht in de moleculaire verschillen die bestaan tussen patiënten, mogelijkheden voor verder onderzoek naar de biologische processen die bijdragen aan het ontstaan van de ziekte en het opsporen van nieuwe targets voor therapeutische interventie.

Genomics om diagnose te verbeteren

Patiënten met gewrichtspijn en die positief zijn voor de autoantistoffen anti-citrulline proteïne (ACPA) en/of reumafactor (RF) hebben een verhoogd risico op het ontwikkelen van reumatoïde artritis. In **hoofdstuk 2** wilden we de mogelijke veranderingen in het immuunsysteem opsporen in een preklinische fase van de ziekte om zo de oorzaak van de ziekte beter te begrijpen en om additionele biomarkers te identificeren met als uiteindelijk doel preventieve medicatie op maat. Patiënten met gewrichtspijn en positief voor de antistoffen ACPA en/of RF bleken een significant verschillend genexpressie profiel te hebben vergeleken met gezonde controles. Dit profiel wordt gekenmerkt door een verhoogde expressie van genen betrokken bij het aangeboren afweermechanisme zoals type I IFN geïnduceerde genen. Gedetailleerde analyse onthulde de aanwezigheid van verschillende subgroepen van patiënten op basis van verschillen in genexpressie profielen wat indicatief is voor interindividuele verschillen in activiteit van het immuun systeem. Sommige preklinische patiënten bleken al een genexpressie profiel te hebben die vergelijkbaar is met die van RA patiënten. Binnen 2 jaar werd bij een aantal preklinische patiënten de diagnose artritis vastgesteld. In dit hoofdstuk beschrijven we een genexpressie profiel dat kan bijdragen aan het voorspellen van artritis in preklinische patiënten.

Genomics voor subclassificatie van patiënten

RA en MS zijn beide heterogene ziektes wat gekenmerkt worden door de variabiliteit in symptomen en respons op medicijnen. In **hoofdstuk 3** hebben we onderzocht of deze heterogeniteit weerspiegeld is in het genexpressie profiel van bloedcellen en

aangetaste weefsels. Zowel de genexpressie profielen in bloedcellen van RA als die van MS patiënten waren aanzienlijk verschillend van die van gezonde individuen en vertoonden een duidelijke verhoging in activiteit van genen betrokken bij de immune respons. Bovendien waren de genexpressie profielen verschillend tussen de patiënten onderling op basis waarvan verschillende subgroepen konden worden onderscheiden. In beide ziektes vonden wij in een subgroep van de patiënten een verhoogde activiteit van type I IFN geïnduceerde genen. Omdat deze genen voornamelijk worden geactiveerd door infectie met pathogenen, zoals virussen, hebben we de genexpressie profielen van patiënten vergeleken met die van virus geïnfecteerde makaken (aapsoort). Een subgroep van de patiënten bleek een genexpressie profiel te hebben die vergelijkbaar is met die van de geïnfecteerde makaken. Dit laatste suggereert dat een virale infectie gerelateerd zou kunnen zijn aan de ontwikkeling of progressie van auto-immuniteit, maar aangezien deze genen ook door lichaamseigen factoren zouden kunnen worden aangezet is nader onderzoek nodig naar de rol van type I IFN genen in MS en RA.
Eerder onderzoek in synoviaal weefsel van RA patiënten heeft laten zien dat er verschillende weefseltypen bestaan. De heterogeniteit van het weefsel suggereert dat verschillende pathogene mechanismen kunnen leiden tot weefselschade die mogelijk verschillend behandeld moeten worden. In *hoofdstuk 3.3* hebben we het genexpressie profiel bepaald van synoviaal weefsel van RA patiënten en deze vergeleken met het bloedprofiel. Wederom konden we op basis van genexpressie profielen twee verschillende typen RA weefsels onderscheiden. Bovendien was de subclassificatie op basis van genexpressie identiek aan subclassificatie op basis van de gemeten infiltratie van ontstekingscellen. Hoewel het weefsel zich duidelijk onderscheid in twee typen, het genexpressie profiel in bloedcellen was vergelijkbaar tussen de twee weefsel typen. De ontstekingsmarker C-reactief proteïne (CRP), de bezinkingssnelheid van erytrocyten (BSE) en de hoeveelheid bloedplaatjes bleken de beste perifere bloedmarkers die de weefselclassificatie weerspiegelen.

Farmacogenomics
Uit klinisch onderzoek weten we dat de behandeling van zowel RA als MS patiënten niet voor iedere patiënt gunstig uitpakt. Het gebrek aan een positieve respons op de therapie suggereert dat sommige patiënten resistent zijn of een ongewenste immuunreactie hebben op de voorgeschreven behandeling. De fysiologische effecten die een medicijn teweeg brengt in ons lichaam noemt men de farmacodynamiek en deze kan verschillend zijn tussen patiënten wat kan leiden tot interindividuele verschillen in klinische respons op behandeling. In **hoofdstuk 4** hebben we de veranderingen in het genexpressie profiel onderzocht die geïnduceerd wordt door therapie. Dit soort onderzoek wordt ook wel "farmacogenomics" genoemd.

Eén van de meest gebruikte behandeling van RA is het remmen van TNF m.b.v. zgn. biologicals zoals Infliximab. Ondanks dat deze middelen in het algemeen doeltreffend

zijn, blijkt dat bij een deel van de patiënten deze medicijnen onvoldoende helpen. Bovendien is de exacte werking van het medicijn nog niet geheel duidelijk. In *hoofdstuk 4.1* hadden wij als doel het bepalen van de farmacogenomic respons op anti-TNF behandeling in RA met behulp van genexpressie analyse van perifere bloedcellen. De genexpressie profielen voorafgaand aan therapie waren significant verschillend vergeleken met die van na therapie. De genen met een verlaagd genexpressie profiel vertegenwoordigen remming van verscheidene immune-gerelateerde biologische processen. Opmerkelijk is dat de door therapie geïnduceerde veranderingen in genexpressie niveaus vergelijkbaar is tussen alle patiënten ongeacht hun klinische respons. Dit duidt aan dat alle RA patiënten een actief TNF respons programma bezitten voorafgaand aan therapie. Aangezien niet alle patiënten goed reageren op behandeling, ondanks de aanwezigheid van een actief TNF respons programma, moeten andere factoren een belangrijke rol spelen bij de klinische respons op therapie.

In *hoofdstuk 4.2* beschrijven we de interactie tussen TNF and type I IFN geïnduceerde genen door het effect van TNF remmers te onderzoeken op het type I IFN genexpressie profiel. De verandering in IFN genexpressie niveaus na TNF blokkade was verschillend tussen patiënten. Sommige patiënten vertoonde een verhoging van het IFN genexpressie profiel terwijl andere patiënten juist een verlaging hadden na TNF remming. Het bleek dat de patiënten met een verhoging van het IFN profiel een slechtere klinische respons hadden op de behandeling.

Type I IFNβ is een frequent toegepaste behandeling om het ziekteproces van MS te remmen. Helaas is de therapie geassocieerd met bijwerkingen en laat niet iedere patiënt een klinisch effect zien. Bovendien is de behandeling kostbaar met slechts een gedeeltelijk klinisch effect waarvan het werkingsmechanisme nog onverklaarbaar is. In *hoofdstuk 4.3* beschrijven we de farmacogenomic response in bloedcellen van MS patiënten na behandeling met type I IFNβ. De farmacogenomic response was gerelateerd aan de activiteit van de type I IFN genen voorafgaand aan therapie. De resultaten in dit hoofdstuk laten zien dat de activiteit van IFN geïnduceerde genen in bloedcellen van MS patiënten voorspellend zou kunnen zijn voor de respons op IFNβ.

Conclusie

De resultaten beschreven in dit proefschrift verschaffen nieuwe inzichten in de relatie tussen genexpressie profielen en de heterogeniteit van chronische ziektes zoals RA en MS. Samenvattend kunnen we concluderen dat het gepresenteerde genomicsonderzoek een belangrijke bijdrage kan leveren aan de ontwikkeling van biomarkers die onderscheid maken tussen verschillende vormen van de ziektes MS en RA en die in de toekomst mogelijk gebruikt kunnen worden voor vroege diagnose, prognose en therapie respons op individueel niveau.

Acknowledgements - Dankwoord

Toen was het plotseling af, op dit meest gelezen stukje tekst na… Promoveren doe je niet alleen, maar met behulp van een grote groep mensen. Om niemand te vergeten wil ik beginnen met iedereen te bedanken die op enige wijze betrokken is geweest bij de totstandkoming van dit proefschrift. En dat beperkt zich zeker niet tot alleen de mensen die betrokken zijn geweest bij de daadwerkelijke inhoud ervan, maar ook mijn vrienden, collega's, familieleden, kennissen, etc. wil ik hartelijk danken voor alle steun en gezelligheid tijdens mijn promotieonderzoek. Een aantal personen wil ik echter nog even apart noemen.

Allereerst wil ik graag mijn promotor Cor Verweij bedanken voor de kans die hij mij gegeven heeft om mijn eigen promotieonderzoek te starten. Ook ben ik dankbaar voor alle vrijheid die je me hebt gegeven, om zo mezelf te ontwikkelen tot een zelfstandig wetenschappelijke onderzoekster. Ik heb veel van je geleerd en zal je nooit vergeten. Veel succes met je nieuwe bedrijf Preselect Diagnostics.
Mijn copromotor Tineke van der Pouw Kraan. Jij hebt me wegwijs gemaakt in de complexe wereld van microarray analyses. Hier ben ik je eeuwig dankbaar voor. Ook onze congresuitjes, die jij altijd perfect vastlegde op de foto, vond ik altijd super gezellig.

Ik ben de leden van de beoordelingscommissie bestaande uit Prof. Dr. R. Mebius, Dr. L. Van 't Veer, Prof. Dr. B.A.C. Dijkmans, Prof. Dr. P.P. Tak, Prof. Dr. C.H. Polman, Prof. Dr. L.A. Aarden en Prof. Dr. G.A. Meijer zeer erkentelijk voor de aandacht die ze hebben geschonken aan mijn proefschrift.

Ben, Wouter, Gertjan en Dirkjan wil ik bedanken voor de prettige samenwerking tijdens het onderzoek voor hoofdstuk 2. Het is een mooi hoofdstuk geworden waar jullie bijdrage onmisbaar is geweest.

En dan nu mijn allerliefste paranimfen Trieneke en Ellen. Samen hebben we lief en leed gedeeld, dat blijkbaar hoort bij het promoveren. Samen hebben we gezorgd voor heel wat gezelligheid die zo belangrijk is naast het harde werken. Ik mis de DE-tjes en theetjes nog steeds. Ellen, ondanks dat we een behoorlijk verschillend onderwerp hadden konden we altijd brainstormen over resultaten en het opschrijven daarvan. Ook was het altijd weer fijn en gezellig om na een avondje stappen bij jou en je konijn te kunnen logeren. Trieneke, als paarse miepen hebben we heel wat af gelachen. Ik heb ook veel geleerd van je onvermoeibare correcties op mijn taalgebruik, er zullen in dit dankwoord ook nog wel wat spelfouten staan, haha, ik krijg ze vast nog wel van je te horen. Heel veel succes met jouw laatste loodjes!

De paarse groep van Cor en Tineke was een nieuwe groep bij de afdeling MCBI die in het prille begin werd gevormd door Josefien, Pia, Nicolette en ik. Pia en Nicolette, het is al weer een hele tijd terug maar ik weet nog goed hoe we als de 3 musketiers samen het labwereldje aan het verkennen waren en op zoek waren naar de invulling van ons project. Nicolette, wat hebben we veel emmers sop gesjouwd en labs schoongemaakt zodat ze geschikt werden voor ons ultraschone RNA werk. Ik vond het jammer dat je na 2 jaar wegging, maar ik ben blij dat je daarna een nieuwe levensweg hebt gevonden en jouw schaterlach blijft me voor altijd bij. Pia, ik heb altijd respect gehad voor jouw doorzettingsvermogen en je literatuurkennis. Ook de stap om je verdere carrière door te brengen in Boston vind ik bewonderenswaardig. Josefien, bedankt voor al je praktische hulp bij mijn experimenten vooral voor de enorme aantallen RNA isolaties die je hebt gedaan en de stimulatieproefjes. Ik kon jou altijd vragen om nog even dit of dat te doen, bedankt! François, al snel kwam jij onze groep versterken. Jouw nuchtere kijk op het leven en natuurlijk de altijd lekker geurende lelies hebben zeker een positieve bijdrage geleverd aan dit boekje. Daarnaast heb je natuurlijk erg veel experimenten en arrays gedaan waar een groot deel van de resultaten van dit proefschrift op gebaseerd zijn. Ontzettend bedankt daarvoor! Anna, jij werd de eerste postdoc in onze groep en ik keek altijd op naar jouw manier van data presenteren en je kennis van epigenetica. Bert, ook jij hebt een korte tijd deel uitgemaakt van onze groep. Bedankt voor je behulpzaamheid en bitterballen op vrijdag. Saskia, jij bent mijn MS project gaan voortzetten en samen hebben we een mooie paper gepubliceerd. Ik wil je bedanken voor de leuke samenwerking en ik wens je veel succes bij de afronding van jouw boekje. Alex, thanks for the Spanish wines en ham, they were always delicious. Good luck with your PhD project in Barcelona! Carina, bij jou kon ik altijd terecht voor een praatje. Jouw opgewektheid en vrolijkheid was altijd fijn en ik heb veel geleerd over zwanger zijn en baby's ;-) Ook alle andere paarse collega's wil ik bedanken voor de goede samenwerking: Auke, Marjolein, Tineke H en Fran. Marianne, ook jij hebt een essentiële (microarray) bijdrage geleverd aan hoofdstuk 4.3, bedankt! Mijn studenten Evert-Jan, John en Suzanne wil ik ook bedanken voor hun inzet tijdens hun stage.

Tijdens mijn loopbaan aan de VU heb ik heel wat kamergenootjes zien komen en gaan en heb ik zelf ook in verschillende (4!) kamers mijn tijd doorgebracht. De laatste maanden van mijn promotieonderzoek verhuisde onze groep naar de afdeling Pathologie in het CCA gebouw. Mijn 1e kamer op de afdeling MCBI was de AIO kamer op de G/H vleugel die ik heb gedeeld met Nicolette, Pia, Klaas, Patricia, Arnold, Stella en Irene. Voor ons allen was dit de start van onze wetenschappelijke carrière en het was dan ook een spannende en leuke tijd. Vervolgens heb ik een tijdje op de analistenkamer mogen doorbrengen met Cees, Edith, Suzanne en Chantal (ik hoop dat ik niemand vergeet…). Jullie werkten al langer op de afdeling en ik heb veel geleerd van jullie (ik vroeg jullie waarschijnlijk de oren van het lijf).

De MCBI groeide en groeide waardoor onze groep moest verhuizen naar het lab en de kamers in de B/C vleugel. Ik ben heel blij dat ik mee mocht denken over de renovatie en inrichting van deze labs. Ik heb tijdens deze periode niet alleen veel geleerd over labstructuur en logistiek, maar ook veel over communicatie en omgang met collega's. Hiervoor wil ik de afdeling en dan in het bijzonder Christien, Ed en Bert hartelijk danken. In de B/C vleugel werd een nieuwe AIO kamer ingericht die al snel de bijnaam "miepenkamer" kreeg. Dit kwam door de fantastische sfeer (met droppot en koekjes) in deze kamer, maar er werd ook heel hard gewerkt. Nicolette en Pia verhuisde met mij mee en de kamer werd later gedeeld met Ellis, Ellen, Trieneke, Marieke, Caroline, Jeroen en Saskia. Alle miepen wil ik bedanken voor de super gezellige tijd, de etentjes, de miepen-drinken-bier-avonden en ons weekendje skiën in België bij Caroline thuis. Het was een mooie onvergetelijke tijd! Succes allemaal, en dat we nog maar vaak miepenuitjes mogen organiseren.

Ik heb een gelukkige, leuke en leerzame tijd gehad bij de afdeling MCBI. De retraites, labuitjes, kerstborrels, AIO weekendjes, de NVVI borrelavonden, B/C-vleugel-vrijdag-koekjes gevolgd door bier en bitterballen. Het was top! Christien, bedankt voor het laatste zetje ;-)

Voor mijn MS gerelateerde onderzoek heb ik veel samengewerkt met de afdeling Neurologie. Lisa van Winsen, samen zijn we begonnen met het verzamelen van PAXgene buizen van MS patiënten om zo een biobank op te zetten voor moleculair onderzoek. Vele buizen zijn verzameld (met uiteraard de hulp van de poli verpleging) en Jessica en Jolijn hebben ook zeker hieraan bijgedragen. Laura en Chris, ik wil jullie bedanken voor jullie onmisbare input en hulp bij hoofdstuk 4.3. Chris, bedankt voor je begeleiding tijdens mijn MS onderzoek en je altijd concrete en behulpzame input. Naast de afdeling Neurologie wil ik ook het MS centrum en de blauwe MCBI groep van Christien hartelijk danken voor alle wetenschappelijke discussies evenals de altijd gezellige MS dagen, MS uitjes en MS congressen.

De hele KIR groep van het AMC ben ik zeer erkentelijk voor het verzamelen van patiënten materiaal en de gezelligheid tijdens congresuitjes zoals de EWRR en ACR. Carla, jouw bijdrage is essentieel geweest voor de verwezenlijking van dit proefschrift. Ook al was het niet altijd eenvoudig om een geschikt moment te vinden om de data te bespreken, samen hebben we een aantal mooie papers geschreven. Bedankt voor je inzet, je optimisme en behulpzaamheid. Marleen en Margriet, de AUTOCURE meetings waren altijd gezellig door jullie aanwezigheid en ik hoop dat we in de toekomst nog vaak samen mogen werken. Paul-Peter, bedankt voor de nieuwe kans die jij me hebt gegeven als teamleider en postdoc in de KIR groep van het AMC. Ik hoop dat we samen nog vele mooie papers mogen schrijven.

Het reumafonds, Stichting MS Research, MSIF, en V-ICI van het VUMC wil ik bedanken voor de financiële bijdrage voor de realisatie van mijn promotieonderzoek. Ook wil ik alle patiënten en vrijwilligers danken die bereid waren mee te werken aan de verschillende onderzoeken. Furthermore, I'm grateful to be part of the AUTOCURE network which provides opportunities for collaboration at an international level.

Dr. Brian Haab and group, thanks for giving me the opportunity to work in your proteomics lab in Michigan USA, where I was trained in Ab-array technology. Beverly, thank you so much for all the great times we spent together in Grand Rapids and for showing me around. You are a wonderful and admirable woman!

Naast het werk is het natuurlijk belangrijk om ook af en toe te kunnen ontspannen en even helemaal ergens anders aan te denken dan aan onderzoek. Turnvereniging St. Mauritius is wat dat betreft mijn 2e werkplek en een soort thuiskomen. Alle leiding, het bestuur, de vele vrijwilligers, de leden en natuurlijk de turn(st)ers, bedankt! Ook bij de dansgroep van Jennifer kan ik wekelijks even lekker uit mijn dak gaan.
Naast sportief ontspannen is het fijn om deel uit te maken van een grote groep lieve vrienden. Bedankt voor jullie interesse, steun en het gezellig samenzijn. Met name, Gertjan, Ina, Afke, Sander, Tineke, Evert-Jan, Ellen, Ramon, Diana, Dennis, Esther, Brenda, Amy, Fons, Marijke, Michel, Marijke, John, Joost, Fieke, Marga en Cees.

Familie Prins, van Baarsen, Weijers, Alexander en ma, eindelijk kan ik jullie het proefschrift overhandigen waar ik de laatste jaren zo druk mee bezig ben geweest. Jullie vertrouwen, zorg en belangstelling is ontroerend, bedankt!

Lieve Marcel, grote broer, het doet pijn dat jij de eindfase van mijn promotieonderzoek niet mag meemaken. Je snapte nooit helemaal wat ik nou eigenlijk zo'n hele dag uitvoerde in het lab, maar je was er volgens mij wel trots op dat jouw kleine zusje doctor zou worden. Ik heb mijn proefschrift aan jou opgedragen, rust zacht.

Liefste Bart, wat zal jij ook blij zijn dat het af is. Eindelijk die papieren bende van de bank, de tafel, de studeerkamer etc ;-) Vele malen heb je mijn geklaag moeten aanhoren als iets weer eens niet ging zoals ik dat gepland had. Je hebt me altijd gesteund, je bent mijn alles. Ik hou van je!

Curriculum Vitae

Lisa van Baarsen werd geboren op 7 augustus 1980 in Volendam. Al tijdens de middelbare school werd ze gefascineerd door de experimenten die uitgevoerd werden tijdens de scheikunde, biologie en natuurkunde lessen. Zodoende begon zij na het behalen van haar HAVO diploma in 1997 aan het Werenfridus te Hoorn, met de studie medische biologie van het Hoger Laboratorium Onderwijs in Amsterdam waarvoor ze in 2001 haar diploma ontving. Tijdens haar stage bij de afdeling Celbiologie in het AMC besloot ze dat ze zelfstandig wetenschappelijk onderzoek wilde uitvoeren en ging op zoek naar een promotieplaats. Zo kwam ze terecht bij Prof. Dr. Cor Verweij op de afdeling Moleculaire Celbiologie en Immunologie aan het VU Medisch Centrum waar ze kon beginnen als research analist met uitzicht op een eigen promotieproject. Eind 2004 startte ze met haar promotieonderzoek naar genexpressie profielen in multiple sclerose en reumatoïde artritis onder supervisie van Cor Verweij en Tineke van der Pouw Kraan. In 2006 was ze te gast in het proteomics laboratorium van Dr. Brian Haab in Michigan USA. De resultaten van haar onderzoek staan beschreven in dit proefschrift. Het onderzoek heeft o.a. geleid tot drie patentaanvragen die de basis zijn voor het bedrijf Preselect Diagnostics B.V. opgericht door Cor Verweij. Momenteel werkt zij als postdoc en teamleider bij de Klinische Immunologie & Reumatologie (o.l.v. Prof. P.P. Tak) in het AMC.

"De opdracht is niet zozeer om te zien wat niemand hiervoor ooit heeft waargenomen; maar na te denken over datgene wat men allemaal kan waarnemen, maar waar tot nu toe nog niet goed over is nagedacht"

Nobelprijswinnaar, kwantumfysicus Erwin Schrödinger (1887-1961)
Uit: Eindeloos bewustzijn, Pim van Lommel

Publication list

L.G.M. van Baarsen, C.L. Bos, T.C.T.M van der Pouw Kraan, C. L. Verweij. Transcription profiling of rheumatic diseases. Arthritis Research & Therapy 2009;11(1):207

C.L. Bos, L.G.M. van Baarsen, T.C.G. Timmer, M.J. Overbeek, N.M. Basoski, F. Rustenburg, J.M.C Baggen, H.J. Thiesen, B.A.C. Dijkmans, T.C.T.M van der Pouw Kraan, A.E. Voskuyl , C.L. Verweij. Molecular subtypes of systemic sclerosis in association with anti-centromere antibodies and digital ulcers. Genes & Immunity 2009 Apr;10(3):210-8

S. Vosslamber, L.G.M. van Baarsen, C.L. Verweij. Pharmacogenomics of IFNb in multiple sclerosis: Towards a personalized medicine approach. Pharmacogenomics 2009 Jan;10(1):97-108

L.G.M.van Baarsen*, S. Vosslamber*, M. Tijssen, J.M. Baggen, L.F. van der Voort, J. Killestein, T.C.T.M. van der Pouw Kraan, C.H. Polman, C.L. Verweij. **Pharmacogenomics of interferon-beta therapy in multiple sclerosis: baseline IFN signature** determines pharmacological differences between patients. PLoS ONE 2008 Apr 2;3(4):e1927

T.C.T.M. van der Pouw Kraan*, C. A. Wijbrandts*, L.G.M. van Baarsen*, F. Rustenburg, J.M. Baggen, C.L. Verweij, P.P. Tak. Responsiveness to anti-TNFα therapy is related to pre-treatment tissue inflammation levels in rheumatoid arthritis patients. Annals of the Rheumatic Diseases 2008 Apr;67(4):563-6

T.C.T.M. van der Pouw Kraan, L.G.M. van Baarsen, F.Rustenburg, B.Baltus, M.Fero, C.L. Verweij. Gene expression profiling in rheumatology. Book chapter in: Methods in Molecular Medicine (ed. Andrew Cope), Arthritis Research 2007;136:305-27

T.C.T.M. van der Pouw Kraan*, L.G.M. van Baarsen*, C.A. Wijbrandts, A.E. Voskuyl, F. Rustenburg, J.M. Baggen, B.A.C. Dijkmans, P.P. Tak, C.L. Verweij. Expression of a pathogen-response program in peripheral blood cells defines a subgroup of rheumatoid arthritis patients. Genes & Immunity 2008 Jan;9(1):16-22

T.C.T.M. van der Pouw Kraan*, C.A. Wijbrandts*, L.G.M. van Baarsen, A.E. Voskuyl, F. Rustenburg, J.M. Baggen, S.M. Ibrahim, M. Fero, B.A.C. Dijkmans, P.P. Tak, C.L. Verweij. Rheumatoid Arthritis subtypes identified by genomic profiling of peripheral blood cells: Assignment of a type I interferon signature in a subpopulation of patients. Annals of the Rheumatic Diseases 2007 Aug;66(8):1008-14

M.C. Lebre, A.M. van der Aar, L.G.M. van Baarsen, T.M. van Capel, J.H. Schuitemaker, M.L. Kapsenberg, E.C. de Jong. Human Keratinocytes Express Functional Toll-Like Receptor 3, 4, 5, and 9. Journal of Investigative Dermatology 2007 Feb;127(2):331-41

L.G.M. van Baarsen, T.C.T.M. van der Pouw-Kraan, C.L. Verweij. A genomic view of subtypes in rheumatoid arthritis: Towards personalized medicine. Future Rheumatology 2006 June Vol. 1, No. 3, Pages 311-322

L.G.M. van Baarsen, T.C.T.M. van der Pouw Kraan, J.J. Kragt, J.M. Baggen, F. Rustenburg, T. Hooper, J.F. Meilof, M.J. Fero, C.D. Dijkstra, C.H. Polman, C.L. Verweij. A subtype of multiple sclerosis defined by an activated immune defense program. Genes & Immunity 2006 Sep;7(6):522-31

P.V. Kasperkovitz, T.C. Timmer, T.J. Smeets, N.L. Verbeet, P.P. Tak, L.G.M. van Baarsen, B.Baltus, T.W. Huizinga, E. Pieterman, M. Fero, G.S. Firestein, T.C.T.M. van der Pouw Kraan, C.L. Verweij. Fibroblast-like synoviocytes derived from patients with rheumatoid arthritis show the imprint of synovial tissue heterogeneity: evidence of a link between an increased myofibroblast-like phenotype and high-inflammation synovitis. Arthritis & Rheumatism 2005 Feb;52(2):430-41

J.P. Bayley, J.G. van Rietschoten, A.M. Bakker, L.G.M. van Baarsen, E.L. Kaijzel, E.A. Wierenga, T.C.T.M. van der Pouw Kraan, T.W. Huizinga, C.L. Verweij. Allele-specific expression of the IL-1 alpha gene in human CD4+ T cell clones. The Journal of Immunology 2003 Sep 1;171(5):2349-53

Submitted for publication:

T. Cantaert*, L.G.M. van Baarsen*, C.A. Wijbrandts, M.G. van de Sande, T.C.T.M. van der Pouw Kraan, C.L. Verweij, P.P. Tak, D.L. Baeten. Type I interferons have no major influence on humoral autoimmunity in rheumatoid arthritis

L.G.M. van Baarsen, C.A. Wijbrandts, T.C.G. Timmer, T.C.T.M. van der Pouw Kraan, P.P. Tak, C.L. Verweij. Synovial tissue heterogeneity in rheumatoid arthritis in relationship to disease activity and biomarkers in peripheral blood.

L.G.M. van Baarsen*, C.A. Wijbrandts*, D.M. Gerlag, F. Rustenburg, T.C.T.M. van der Pouw Kraan, B.A.C. Dijkmans, P.P. Tak, C.L. Verweij. Pharmacogenomics of infliximab treatment using peripheral blood cells of patients with rheumatoid arthritis.

L.G.M. van Baarsen, C.A. Wijbrandts, F. Rustenburg, T. Cantaert, T.C.T.M. van der Pouw Kraan, D.L. Baeten, B.A.C. Dijkmans, P.P. Tak, C.L. Verweij. Regulation of IFN response gene activity during infliximab treatment in rheumatoid arthritis is associated with clinical response to treatment

L.G.M. van Baarsen, W.H. Bos, F. Rustenburg, T.C.T.M. van der Pouw Kraan, G.J. Wolbink, B.A.C. Dijkmans, D. van Schaardenburg, C.L. Verweij. Gene expression profiling in autoantibody positive arthralgia patients predicts development of arthritis.

* Authors contributed equally

Patents:

C.L. Verweij, L.G.M. van Baarsen, C.H. Polman, S. Vosslamber. Means and methods for classifying samples of multiple sclerosis patients

Patent application filed 2007

C.L. Verweij, L.G.M. van Baarsen, W. Bos, D. van Schaardenburg. Preclinical biomarkers for predicting the development of chronic autoimmune diseases.

Patent application filed 2008

C.L. Verweij, L.G.M. van Baarsen. Predicting clinical respons to treatment with a soluble TNF-antagonist or TNF, or a TNF receptor agonist.

Patent application filed 2008